INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO
MESTRADO EM ENGENHARIA INFORMÁTICA

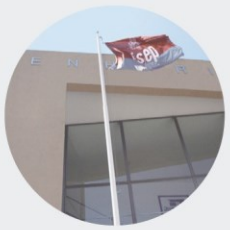



Curadoria Digital em dados: proposta de um modelo de preservação digital em dados na área da saúde

ANA PAULA RABELO DE FREITAS
Novembro de 2024

POLITÉCNICO
DO PORTO

INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO
MESTRADO EM ENGENHARIA INFORMÁTICA




Curadoria Digital em dados: proposta de um modelo de preservação digital em dados na área da saúde

ANA PAULA RABELO DE FREITAS

Novembro 2024

POLITÉCNICO
DO PORTO

ISEP INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

Curadoria Digital em dados: proposta de um modelo de preservação digital em dados na área da saúde

Ana Paula Rabelo de Freitas

Dissertação para obtenção do Grau de Mestre em Engenharia Informática, Área de Especialização em Engenharia de Software

Coorientador: Prof. Dr. Nuno Escudeiro

Coorientadora: Prof.ª Dr.ª Jeane Silva Ferreira

Porto, Novembro 2024

Declaração de Integridade

Declaro ter conduzido este trabalho acadêmico com integridade.

Não plagiei ou apliquei qualquer forma de uso indevido de informações ou falsificação de resultados ao longo do processo que levou à sua elaboração.

Portanto, o trabalho apresentado neste documento é original e de minha autoria, não tendo sido utilizado anteriormente para nenhum outro fim.

Declaro ainda que tenho pleno conhecimento do Código de Conduta Ética do P. PORTO.

ISEP, Porto, 20 de Novembro de 2023

Resumo

Este trabalho tem como objetivo propor um modelo de preservação de dados digitais na área da saúde, com foco na preservação dos registros médicos. A gestão eficaz de dados de saúde tornou-se uma prioridade devido ao avanço tecnológico e à importância do acesso seguro e preservação de informações médicas. No entanto, o setor ainda enfrenta desafios significativos, como a rápida obsolescência tecnológica e as regulamentações de sigilo médico. Diante da diversidade dos dados e das exigências legais, é crucial desenvolver abordagens proativas para garantir a integridade e disponibilidade dessas informações ao longo do tempo. A pesquisa adotará uma abordagem indutiva qualitativa, utilizando a observação, coleta e análise de dados para identificar princípios, normas e diretrizes que orientarão o desenvolvimento do modelo proposto. Além disso, buscará compreender as relações existentes na curadoria e preservação digital na área da saúde, a fim de oferecer um plano de ação fundamentado e eficaz para a preservação de informações médicas na era digital.

Palavras-chave: Preservação digital, Saúde digital, Curadoria digital, Governança de dados.

Abstract

This work aims to propose a model for the preservation of digital data in the healthcare sector, focusing specifically on the preservation of medical records. Effective management of health data has become a priority due to technological advances and the importance of ensuring secure access and preservation of medical information. However, the sector still faces significant challenges, such as rapid technological obsolescence and strict regulations concerning medical confidentiality. Given the diversity of data and legal requirements, it is crucial to develop proactive approaches to ensure the integrity and availability of this information over time. The research will adopt an inductive qualitative approach, employing observation, data collection, and data analysis to identify the principles, standards, and guidelines that will guide the development of the proposed model. Additionally, it seeks to understand the existing relationships in digital curation and preservation within the healthcare domain, with the aim of providing a well-founded and effective action plan for preserving medical information in the digital age.

Índice

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e siglas

ABCD	Access to Biological Collection Data.
AIP	Archival Information Package.
API	Application Programming Interface.
APS	Atenção Primária à Saúde.
AVM	Astronomy Visualization Metadata Standard.
BRAPCI	Base de Dados em Ciência da Informação.
BSEN	Biblioteca Setorial de Enfermagem e Nutrição.
CBHPM	Classificação brasileira hierarquizada de procedimentos médicos.
CCSDS	Consultative Committee for Space Data Systems.
CETIC	Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação.
CFM	Conselho Federal de Medicina.
CIAP-2	Classificação Internacional de Atenção Primária.
CID	Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde.
CONARQ	Conselho Nacional de Arquivos.
COPTR	Community Owned Digital Preservation Tool Registry.
DataONE	Data Observation Network for Earth.
DATASUS	Departamento de Informática do SUS.
DC	DublinCore.
DCC	Digital Curation Centre.
DCU	Digital Curation Unit.
DDI	Data Documentation Initiative.

DGI	Data Governance Institute.
DICOM	Digital Imaging and Communications in Medicine.
DMBOK	Data Management Body of Knowledge.
DPC	Digital Preservation Coalition.
DwC	Darwin Core.
EBSERH	Empresa Brasileira de Serviços Hospitalares.
ENAP	Escola Nacional de Administração Pública.
ESD28	Estratégia de Saúde Digital para 2028.
FAIR	Findable, Accessible, Interoperable, Reusable.
FBN	Fundação Biblioteca Nacional.
FEI	Front End of Innovation.
FHIR	Fast Healthcare Interoperability Resources.
FioCruz	Fundação Oswaldo Cruz.
HL7	Health Level 7.
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia.
ICICT	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde
InterPARES	International Research on Permanent Authentic Records in Electronic Systems.
ISBT	International Society of Blood Transfusion.
ISO	International Organization for Standardization.
LGPD	Lei Geral de Proteção de Dados.
LOCKSS	The Lots of Copies Keeps Stuff Safe.
LOINC	Logical Observation Identifiers Names and Codes.
MS	Ministério da Saúde.

NCD	New Concept Development.
NDSA	National Digital Stewardship Alliance.
NPD	New Product Development.
OAIS	Open Archival Information System.
OMS	Organização Mundial de Saúde.
openEHR	Open Eletronic Health Records.
PDI	Preservation Description Information.
PEC	Prontuário Eletrônico do Cidadão.
PEP	Prontuário Eletrônico do Paciente.
PNIIS	Política Nacional de Informação e Informática em Saúde.
PPDBN	Política de Preservação Digital da Biblioteca Nacional.
PPDig@ES	Política de Preservação Digital do Governo do Estado do Espírito Santo.
PIX	Patient Identifier Cross-Referencing.
QEF	Quantitative Evaluation Framework.
RES	Registros Eletrônico de Saúde.
RGPD	Regulamento Geral sobre a Proteção de Dados.
RNDS	Rede Nacional de Dados em Saúde.
RNP	Rede Nacional de Ensino e Pesquisa.
RTS	Repositório de Terminologias de Saúde.
S-RES	Sistema de Registro Eletrônico de Saúde.
SINPRED	Seminário Internacional de Preservação Digital.
SIP	Submission Information Packages.
SISAB	Saúde para a Atenção Básica.
SNOMED-CT	Systematized Nomenclature of Medicine - Clinical Terms.

SUS	Sistema Único de Saúde.
SWOT	Strengths, Weaknesses, Opportunities e Threats.
TEI	Text Encoding Initiative Guidelines.
TIC	Tecnologias de Informação e Comunicação.
TISS	Troca de Informações na Saúde Suplementar.
UNESCO	United Nations Educational, Scientific and Cultural Organization.
UNIRIO	Universidade Federal do Estado do Rio de Janeiro.
VSD	VideoSaúde-Distribuidora.

1 Introdução

Neste capítulo é feita uma apresentação do contexto em que esta pesquisa se insere e são identificados o problema, os objetivos e a metodologia desta dissertação.

1.1 Contexto

A saúde é reconhecida como um direito inalienável de todos os indivíduos. No Brasil, essa prerrogativa é assegurada pela Constituição da República Federativa de 1988, juntamente com um conjunto de leis que estabelece que a responsabilidade pela promoção das condições essenciais para o pleno exercício do direito à saúde recai sobre o Estado, envolvendo órgãos, instituições da Administração direta e indireta, e Fundações mantidas pelo Poder Público (BRASIL, 1988). Paralelamente, à medida que a área da saúde avança na adoção de tecnologias digitais para melhorar o atendimento médico, o armazenamento e a preservação dos registros de saúde tornam-se questões críticas.

A área da saúde conta com a ampla adoção de ferramentas digitais por médicos e hospitais para a gestão de processos clínicos, administração hospitalar e métodos de diagnóstico. Dessa forma, os pacientes dependem da capacidade dessas organizações em garantir o acesso contínuo aos seus registros ao longo de toda a vida. Isso inclui também todos os meios de diagnóstico, como radiografias, tomografias computadorizadas, eletrocardiogramas e ultrassonografias, que atualmente são produzidos e armazenados em formatos digitais (FERREIRA, M.; SARAIVA, R.; RODRIGUES, E., 2012). A garantia de que as informações médicas dos pacientes permaneçam acessíveis, autênticas e íntegras ao longo do tempo é vital para garantir

a eficácia dos sistemas de saúde, a proteção dos direitos dos indivíduos e da valorização destes dados como ativos digitais[1] estratégicos.

Esses dados digitais de saúde possuem estruturas diversificadas e muitas das informações relacionadas a eles estão protegidas pelo sigilo médico, com acesso permitido com base na classificação do conteúdo, conforme definido na Constituição Federal (BRASIL, 1988), no Código de Ética Médica e resoluções do Conselho Federal de Medicina (CFM, 2019), na Lei nº 12.527 – Lei de Acesso à Informação (BRASIL, 2011a), e na Lei nº 13.709 – Lei Geral de Proteção de Dados Pessoais (LGPD) (BRASIL, 2018a).

Diante desse cenário, surge uma série de desafios relacionados à preservação desses dados. Para tanto, as organizações precisam equilibrar a gestão de dados a curto e longo prazo, desenvolvendo estratégias que envolvem princípios, políticas e processos específicos para assegurar o controle e o valor desses ativos ao longo de seus ciclos de vida (DAMA, 2017). Nesse contexto, as estratégias de preservação digital têm o objetivo de proteger, cuidar e garantir a longevidade de objetos digitais, para que permaneçam interpretáveis mesmo diante das mudanças tecnológicas e organizacionais (WEBB, 2003).

1.2 Problema

No Brasil, a Lei nº 13.787 (BRASIL, 2018b) trouxe regras claras e diretrizes específicas para a digitalização e o uso de sistemas informatizados para a guarda, o armazenamento e o manuseio de prontuários médicos[2]. A lei determina o prazo mínimo de 20 anos para a guarda destes registros tanto em formato digital quanto

---

[1] Um ativo digital é um conteúdo armazenado em um formato de arquivo digital. Ativos digitais incluem fotos, vídeos, imagens, gráficos, fontes, áudios, apresentações, documentos de texto, e outros conteúdos.

[2] O prontuário é um documento legal que todo paciente precisa ter para ser atendido. Neste documento consta a história de atendimento do paciente no hospital, as consultas, solicitações de exames, cirurgias feitas ou agendadas.

em papel. Já o art. 7º da Resolução nº 1.821/2007 (CFM, 2007) do Conselho Federal de Medicina (CFM) estabelece a guarda permanente para os prontuários dos pacientes arquivados eletronicamente. A resolução ainda reforça que os dados contidos nos prontuários pertencem ao paciente, mas ficam sob a guarda e total responsabilidade da instituição de saúde que os produziu. Dessa forma, o paciente tem o direito de solicitar cópias ou alterações das informações de seu prontuário a qualquer momento, e é vedado ao médico "negar ao paciente ou a seu representante legal, acesso a seu prontuário, deixar de lhe fornecer cópia quando solicitada, bem como deixar de lhe dar explicações necessárias à sua compreensão, salvo quando ocasionarem riscos ao próprio paciente ou a terceiros" (CFM, 2019).

A preservação adequada dos prontuários médicos vai além do cumprimento da legislação; elas têm um impacto direto na preservação da história médica de pacientes, e são fundamentais para garantia da continuidade dos cuidados e tratamentos. Além disso, os dados de saúde contidos nesses registros são elementos primordiais para a elaboração de políticas públicas baseadas em evidências. A Lei nº 8.080/1990, que regulamenta o Sistema Único de Saúde (SUS), estabelece a necessidade de utilizar dados de saúde para o planejamento, execução e monitoramento das políticas de saúde pública no Brasil (BRASIL, 1990). Esses dados alimentam diversos indicadores de saúde, que são utilizados para monitorar a qualidade e a efetividade dos serviços de saúde prestados à população, além de serem uma ferramenta essencial para a alocação de recursos e a definição de estratégias de atendimento em todo o país.

No entanto, os objetos digitais enfrentam diversas ameaças ao longo de sua existência. Conforme descrito pelo DPC (2015), uma das principais ameaças diz respeito às mídias de armazenamento que podem se deteriorar, se tornar obsoletas, ou até serem acidentalmente deletadas ou destruídas maliciosamente, levando à corrupção de arquivos e, consequentemente, à perda de dados. Além disso, há a ameaça de que os dados percam seu significado ao longo do tempo. Formatos de arquivos e *softwares* utilizados para armazenar e acessar esses dados podem se tornar obsoletos, tornando difícil a recuperação de informações essenciais para a continuidade do tratamento médico ou para a análise de longo prazo dos prontuários. Esse processo de obsolescência tecnológica pode resultar na perda de informações, ou, no mínimo, em interpretações imprecisas ou falhas de dados.

Outro problema é a vulnerabilidade da integridade dos dados. Objetos digitais, por natureza, são suscetíveis a alterações involuntárias ou intencionais: erros de mídia, falhas de *software* ou ações humanas podem corromper arquivos e mesmo pequenas alterações podem comprometer a autenticidade e confiabilidade de um registro médico. Há ainda a ameaça da perda de contexto dos dados ao longo do tempo. Registros médicos digitais podem depender de outros arquivos ou informações contextuais que, se não forem preservados adequadamente, podem dificultar a interpretação correta dos dados no futuro. A falta de metadados ou a perda de informações sobre como os dados foram originalmente utilizados pode criar barreiras à compreensão completa dos registros (DPC, 2015).

Por fim, o crescimento exponencial dos dados gerados supera a capacidade de armazenamento e os recursos financeiros destinados à sua preservação. Isso cria um cenário onde a incapacidade de gerenciar e selecionar os dados mais críticos pode resultar em uma sobrecarga de informações, o que dificulta a gestão eficiente dos objetos digitais e aumenta o risco de perdas de dados.

1.3 Objetivos

Diante destas ameaças, o principal objetivo dessa pesquisa é desenvolver um modelo de preservação digital para bases de dados na área de saúde no contexto brasileiro em conformidade com padrões recomendados.

Para orientar de maneira mais precisa a pesquisa e garantir o alcance do objetivo geral, delinearam-se os seguintes objetivos específicos:

a) Analisar as etapas, padrões e normas para a implantação de um plano de preservação digital;

b) Explorar o panorama atual de diretrizes de preservação digital, a partir de uma revisão da literatura;

c) Identificar e caracterizar as necessidades específicas de preservação digital no contexto das bases de dados de saúde, considerando o contexto brasileiro;

d) Propor um modelo de preservação digital que considere as características das bases de dados de saúde.

1.4 Perguntas norteadoras

Esta pesquisa se insere em um contexto mais amplo de preocupações crescentes com a preservação da informação digital, e visa propor um modelo que leva em consideração os desafios específicos dessa área, incluindo a diversidade de dados, as regulamentações de sigilo médico e as tecnologias em constante evolução. Assim, o tema proposto nesta pesquisa busca estudar a área de preservação digital e os possíveis modelos para esses ativos digitais. Diante do exposto, as seguintes perguntas norteadoras serão consideradas para o desenvolvimento do trabalho:

1. Qual o panorama atual da Preservação Digital no Brasil?
2. Quais as dificuldades de Preservação Digital relacionado à área de aplicação (saúde)?
3. Quais as etapas, padrões e normas para a implantação de um plano de Preservação Digital?
4. Como utilizar os planos e políticas de Preservação Digital já existentes como suporte para criar um modelo compatível com as características de dados na área de saúde?
5. Quais os critérios para definição de um modelo compatível com estes padrões no contexto pré-definido?

1.5 Metodologia de pesquisa

Essa pesquisa tem como objeto de estudo a elaboração de um modelo de curadoria digital, com foco na preservação digital de dados na área de saúde, com o objetivo de abordar os desafios específicos relacionados à crescente quantidade de registros digitais na área da saúde e à importância da preservação da informação médica de pacientes.

O modelo pretendido caracteriza-se como um plano de preservação digital, e reunirá elementos essenciais para a execução eficaz da preservação digital, desde a identificação e contexto inicial até a implementação. Ele definirá passos específicos e estabelecerá uma estrutura para a manutenção contínua e revisões periódicas, assegurando que possa adaptar-se às mudanças no ambiente institucional e tecnológico ao longo do tempo.

1.5.1 Método da pesquisa

Uma revisão de literatura narrativa foi conduzida na primeira fase desta pesquisa e abrangeu uma busca utilizando uma combinação de palavras-chave relacionadas à saúde digital, curadoria digital, preservação digital e governança e gestão de dados. O método utilizado está documentado nos seguintes passos.

1. Questões da pesquisa

   Para direcionar o método adotado, a revisão da literatura foi orientada com base nas perguntas norteadoras definidas na seção 1.4 – Perguntas norteadoras. As perguntas foram adaptadas ao contexto da revisão:

   QP1 – Qual o cenário nacional e internacional nas áreas de Curadoria e Preservação Digital?

   QP2 – Qual o panorama atual de políticas, planos e estratégias de Preservação Digital no contexto brasileiro?

   QP3 – Quais as etapas, padrões e normas para a implantação de um plano eficaz de Preservação Digital?

   A QP1 buscou contextualizar a situação atual e identificar as principais tendências que possam influenciar estratégias futuras nas áreas da curadoria e preservação digital. Já a QP2 direciona o foco para o cenário específico brasileiro, buscando entender o estado atual das diretrizes de Preservação Digital aplicadas no país. A compreensão dessas diretrizes é fundamental para avaliar a eficácia das medidas existentes, identificar oportunidades de aprimoramento e servir como base para o modelo pretendido.

A QP3, por sua vez, complementa as duas primeiras ao buscar identificar as etapas, padrões e normas essenciais para a implementação de um plano eficaz de preservação digital. Essa pergunta busca identificar as diretrizes práticas que possam guiar a definição de estratégias específicas para a gestão de dados na área de saúde, de forma a permitir a aplicação das melhores práticas de preservação digital adaptadas às necessidades desse setor.

2. Fontes de informação

Para a seleção de artigos, livros, relatórios ou teses relevantes que se adequassem às áreas exploradas na revisão, foi utilizado o Google Acadêmico[3], a Base de Dados em Ciência da Informação (BRAPCI)[4] e o *Scientific Electronic Library Online* (SciELO)[5] como principais fontes de pesquisa. Inicialmente, termos como "curadoria digital", "preservação digital" e "saúde digital" foram utilizados como base para a busca. Posteriormente, esses termos foram refinados e direcionados para o contexto específico da tese. Além disso, foram considerados outros trabalhos frequentemente citados como referências em outras publicações, ampliando o escopo da pesquisa e garantindo a inclusão de fontes relevantes e consolidadas.

Durante o processo de pesquisa, foram identificadas revistas e organizações específicas como importantes fontes de conteúdo relacionado às áreas de preservação e curadoria digital, que foram selecionadas como fornecedores de materiais para a presente revisão. Dentre os recursos disponibilizados por

---

[3] O Google Acadêmico, ou Google Scholar, é uma plataforma de pesquisa lançada pelo Google que reúne um acervo de publicações de conteúdo científico provenientes de editoras, sociedades profissionais, repositórios online e universidades. Disponível em: https://scholar.google.com/.

[4] A BRAPCI é uma plataforma digital brasileira dedicada à coleta, preservação e ao acesso de literatura científica na área de Ciência da Informação. Ela abrange uma ampla gama de publicações, incluindo artigos de periódicos, trabalhos de eventos, livros e capítulos de livros, principalmente de fontes brasileiras e américa latina. Disponível em: https://brapci.inf.br.

[5] O SciELO é uma biblioteca digital de livre acesso e um projeto cooperativo de publicação digital de periódicos científicos. Disponível em: https://www.scielo.org/en/.

essas entidades, foram selecionadas publicações que abordam questões cruciais relacionadas às áreas e sua aplicação na saúde.

3. Critérios de seleção

   Inicialmente, foi realizada uma leitura dinâmica dos materiais para avaliar rapidamente sua relevância em relação aos objetivos da pesquisa. Em seguida, quaisquer referências duplicadas foram desconsideradas. Por fim, os critérios de seleção incluíram a exigência de que os artigos estivessem escritos em português ou inglês, fossem de acesso aberto ou disponibilizados gratuitamente, tivessem uma ligação direta com os temas da pesquisa e se relacionasse diretamente à área da Tecnologia da Informação (TI), que foram então utilizados para remover quaisquer fontes que não respondessem diretamente as perguntas de pesquisa ou os objetivos.

1.5.2 Caracterização da pesquisa

O método indutivo, como definido por Lakatos e Marconi (2007, p. 86), se revela a escolha apropriada no que se refere à abordagem para esta pesquisa, pois "envolve o processo de generalização a partir de dados particulares suficientemente constatados" (*apud* PRODANOV; FREITAS, 2013, p. 28). A pesquisa será qualitativa, pois partirá da observação de fatos e fenômenos relacionados à curadoria e preservação digital, buscando compreender as relações existentes para extrair princípios, normas e diretrizes que nortearão a criação do modelo proposto. Por meio da coleta, observação e experimentação, serão investigadas as conexões entre diversos elementos da realidade nacional e internacional. Esse enfoque nos permitirá chegar a conclusões gerais que proporcionarão uma abordagem fundamentada e eficaz para o modelo de preservação da informação médica na era digital.

Esta pesquisa é classificada como aplicada em relação à sua natureza, conforme a definição de Prodanov e Freitas (2013, p. 51), devido ao seu objetivo fundamental de gerar conhecimentos direcionados à solução de problemas específicos na área de saúde. A abordagem aplicada deste estudo busca não apenas compreender os desafios relacionados à crescente quantidade de registros digitais na área da saúde, mas também contribuir ativamente para a preservação de registros médicos.

Sob o ponto de vista dos seus objetivos, essa pesquisa se enquadra principalmente como uma pesquisa exploratória e descritiva, que desempenharão papeis cruciais na fase inicial deste estudo. A pesquisa exploratória visa proporcionar uma compreensão mais aprofundada do problema em questão e dos objetivos definidos. Neste projeto, a exploração envolve a identificação de etapas, padrões, normas e melhores práticas em curadoria digital e sua adaptação para a preservação de dados na área de saúde. Além disso, nesta etapa será realizada uma revisão da literatura que permitirá a definição de um contexto sólido para a pesquisa, identificando oportunidades em relação às diretrizes de preservação existentes, principalmente, no cenário brasileiro. A identificação de estudos de caso e análise dos desafios e necessidades específicas relacionadas à área de saúde também serão aplicadas à pesquisa afim de gerar conhecimentos práticos da área.

A pesquisa descritiva, além da compreensão do contexto da curadoria digital, permitirá registrar e descrever de maneira padronizada as características das práticas de curadoria existentes em outros cenários e adaptá-las à área de saúde. Além disso, possibilitará a ordenação e análise sistemática dos dados encontrados na literatura, identificando relações entre políticas nacionais e internacionais de preservação digital, bem como suas características distintas. No que diz respeito à análise dos desafios e riscos associados aos principais modelos e cases de preservação digital encontrados na literatura, será possível elencar critérios de comparação e descrever os aspectos identificados de forma objetiva, auxiliando na identificação de lacunas e áreas de aprimoramento.

Quanto aos procedimentos técnicos, ou seja, a maneira pela qual coletaremos os dados necessários para a elaboração da pesquisa (PRODANOV; FREITAS, 2013, p. 54), a pesquisa se vale das chamadas fontes de papel, sendo categorizada como pesquisa bibliográfica e documental. A pesquisa bibliográfica faz uso de materiais previamente publicados, abrangendo recursos como livros, revistas, periódicos, artigos científicos, entre outros, para proporcionar uma compreensão sólida e abrangente do tópico de pesquisa em questão (Prodanov e Freitas, 2013, p. 54). A pesquisa documental será empregada de forma complementar à pesquisa bibliográfica, visando enriquecer a obtenção de dados necessários para o desenvolvimento desta investigação.

De acordo com Prodanov e Freitas (2013, p. 55), esse tipo de pesquisa se distingue da bibliográfica por se basear em materiais que não receberam um tratamento analítico ou que podem ser adaptados de acordo com os objetivos da pesquisa. Isso envolve documentos de primeira mão, como documentos oficiais, cartas, relatórios, entre outros, e documentos de segunda mão, como relatórios de pesquisa e tabelas estatísticas que já foram analisados anteriormente. Ambos os tipos de pesquisa, bibliográfica e documental, contribuirão para fornecer uma base sólida para a revisão da literatura, permitindo a compreensão das políticas, planos e estratégias de preservação digital e casos de estudo relevantes, além de permitir o registro, descrição, análise de práticas e para o estabelecimento do contexto da curadoria digital no campo da saúde.

1.6 Estrutura do trabalho

Este trabalho inicia com uma **Introdução** que apresenta um panorama geral dos aspectos principais da saúde digital, destacando a crescente utilização de tecnologias que produzem e armazenam registros médicos em formato digital e a importância da preservação digital nesse contexto. A relevância desse tema é contextualizada ao justificar a necessidade de garantir o acesso seguro, autêntico e contínuo a esses dados, que são essenciais tanto para o tratamento médico dos pacientes quanto para a pesquisa e formulação de políticas de saúde. Em seguida, o problema da pesquisa é exposto, identificando os desafios de preservar informações digitais a longo prazo, considerando ameaças como a obsolescência tecnológica e a vulnerabilidade à perda de dados.

O capítulo inclui também a definição dos objetivos gerais e específicos da pesquisa, com o principal intuito de desenvolver um modelo de preservação digital voltado para o setor de saúde no Brasil. Para orientar o desenvolvimento do trabalho, foram formuladas perguntas norteadoras que exploram questões sobre a preservação digital na área de saúde, os padrões e normas existentes, e como adaptar essas diretrizes para o contexto brasileiro.

No **Estado da Arte**, são discutidos os principais conceitos e modelos de preservação digital, com um enfoque também nas práticas de curadoria digital e governança de

dados. Alguns modelos relevantes são revisados, como o Open Archival Information System (OAIS) e o Digital Curation Lifecycle Model (DCC), explicando suas etapas e importância na preservação de dados digitais. Ao abordar a saúde digital, a revisão explora a necessidade de uma governança de dados bem desenvolvida no setor, dada a sensibilidade dos dados médicos e a importância da interoperabilidade entre sistemas para garantir a acessibilidade e a segurança das informações.

Além dos modelos, a revisão de literatura também examina as regulamentações e normas aplicáveis ao Brasil, como a Lei Geral de Proteção de Dados (LGPD) e a Lei nº 13.787/2018, que estabelece diretrizes para a digitalização e armazenamento de prontuários médicos eletrônicos. Esses aspectos são cruciais para entender o contexto legal e técnico em que o modelo de preservação será aplicado, permitindo a identificação de lacunas para a preservação de dados na saúde.

No capítulo de **Análise de Valor**, a pesquisa se concentra em examinar como o modelo de preservação digital proposto pode oferecer o máximo de valor ao usuário final, com ênfase em um uso eficiente de recursos e minimização de custos, preservando a qualidade e integridade dos dados. Esse processo parte de uma compreensão aprofundada das necessidades e expectativas dos clientes – instituições de saúde, profissionais da área, pacientes e governo – e de como o modelo pode agregar valor ao reduzir os riscos associados à perda de dados críticos e a promover a interoperabilidade entre sistemas de saúde. Por meio do entendimento dos ganhos e sacrifícios envolvidos para o usuário, a pesquisa consegue avaliar não apenas os aspectos técnicos do modelo, mas também o impacto que ele terá no cotidiano das instituições de saúde e na confiabilidade dos dados preservados.

O modelo de preservação digital a ser desenvolvido é estruturado na **Experimentação**, que apresenta as etapas, responsabilidades e recursos necessários para sua implementação. No capítulo **Modelo de Preservação Digital**, o modelo é desenvolvido em várias fases, começando pela identificação e avaliação dos dados médicos quanto à sua importância e sensibilidade. Em seguida, são abordados os processos de ingestão e armazenamento, que garantem a integração dos dados ao sistema de preservação, associando-os a metadados robustos que facilitam o gerenciamento a longo prazo. O modelo também inclui procedimentos de monitoramento contínuo e auditorias para assegurar a integridade dos dados, além de estratégias para garantir o acesso controlado e seguro. Outro aspecto importante é a

definição de métodos para lidar com a obsolescência tecnológica, com orientações para a migração de formatos e compatibilidade com novas tecnologias.

Na **Avaliação**, critérios específicos foram utilizados para medir a qualidade do modelo de preservação digital proposto, incluindo aspectos como eficácia, conformidade com regulamentações, viabilidade econômica e eficiência na disponibilidade e integridade de informações. A avaliação é realizada por meio de uma análise quantitativa, utilizando o modelo de avaliação *Quantitative Evaluation Framework* (QEF), que fornece uma estrutura objetiva para quantificar o desempenho do modelo em relação a cada critério. O modelo permite atribuir pontuações aos diferentes aspectos avaliados, facilitando a comparação entre requisitos e a identificação de áreas que precisam de melhorias. Os resultados obtidos destacam as vantagens e limitações do modelo proposto e permite analisar se o modelo atende aos requisitos pretendidos.

No capítulo de **Objetivos Concretizados**, a pesquisa apresenta uma síntese dos objetivos estabelecidos inicialmente e avalia até que ponto eles foram alcançados ao longo do desenvolvimento do modelo de preservação digital. Esta seção examina o progresso em relação aos objetivos gerais e específicos, detalhando como cada um deles foi abordado, as metodologias aplicadas e os resultados obtidos. Além disso, destaca os principais avanços obtidos com o modelo proposto. Em **Limitações e Trabalhos Futuros**, o estudo reconhece as restrições encontradas ao longo do processo de desenvolvimento e implementação do modelo. Esta seção explora como essas limitações influenciaram os resultados e oferece uma visão sobre os pontos que poderiam ser aprimorados. Além disso, são sugeridas direções para pesquisas futuras.

Por fim, a **Conclusão** oferece uma reflexão sobre o impacto da pesquisa para o campo da preservação digital na saúde, sintetizando os principais achados e contribuições. Este capítulo analisa como o modelo de preservação digital desenvolvido pode auxiliar instituições de saúde na gestão a longo prazo de dados críticos, garantindo sua acessibilidade e integridade, além de promover a segurança e confiabilidade essenciais para o setor. A conclusão destaca também o potencial do modelo para estabelecer uma referência para futuras implementações em outras instituições, reforçando seu valor enquanto ferramenta de preservação e contribuindo para uma base de práticas de preservação digital no contexto da saúde.

2 Estado da Arte

A preservação digital é um desafio crescente em um mundo onde a informação e o conhecimento são cada vez mais digitais. No entanto, torna-se uma tarefa complexa e desafiadora, especialmente devido à rápida obsolescência tecnológica e à dependência de contextos tecnológicos específicos para a interpretação e acessibilidade dos objetos digitais.

A fim de contextualizar e compreender a evolução e o cenário atual das áreas de Curadoria Digital e Preservação Digital, esse capítulo aborda conceitos e metodologias relevantes relacionadas à gestão e conservação de objetos digitais. Além disso, será analisado o panorama atual dos temas na saúde digital no Brasil, com enfoque nas questões técnicas, legais e organizacionais que permeiam a utilização de tecnologias digitais na área da saúde.

2.1 Objetos Digitais

Como define ARELLANO (2008), "objetos digitais são tipos de arquivos encontrados em meio digital, compostos de conjuntos de sequências de bits[6] sobre conteúdos informacionais, metadados e identificadores". Essa definição ressalta a natureza complexa dos objetos digitais, que vão além de simplesmente armazenar informações. Salvar os *bits* de um objeto digital é necessário, porém, é necessário também conhecer os atributos da aplicação na qual ele foi criado e com o qual ele pode ser interpretado. O esforço da preservação é a parte mais longa e a última do ciclo de gerenciamento de objetos digitais. (MÁRDERO ARELLANO, 2008)

Ao montar uma infraestrutura tecnológica para gerenciar estes objetos digitais, as instituições deparam-se com uma série de desafios e decisões estratégicas. O

---

[6] Um dígito binário (*bit*) é a menor unidade de informação armazenada ou transmitida em um computador. Um *bit* tem um valor único: 0 ou 1.

ambiente no qual esses objetos digitais estão inseridos, apresenta vulnerabilidades relacionadas ao desgaste ou falha das mídias; à obsolescência de *hardware* e *software*; falta de manutenção ou expansão da infraestrutura; a falhas humanas; a não aplicação de processos curatoriais ou negligência, que acabam por reduzir o seu tempo de vida útil ou causar a sua perda. (SIEBRA, 2019) Desse modo, é necessário entender a diversidade e a complexidade dos objetos digitais, avaliando fatores como quantidade, tamanho, controle e valor institucional de cada objeto (GALINDO; LA FUENTE, 2015 *apud* CAVALCANTI MOREIRA, 2017).

A aplicação de objetos digitais na saúde é um campo em rápida evolução, com implicações significativas para a preservação de informações vitais e a melhoria dos cuidados aos pacientes. Esses materiais digitais incluem os nato-digitais, que são aqueles criados diretamente em formato digital (ACADEMIA BRASILEIRA DE LETRAS, 2021), como os registros eletrônicos de saúde (RES), imagens de diagnóstico, dados de sensores vestíveis e os documentos digitalizados[7].

Enquanto os objetos nato-digitais nascem já no ambiente digital, os digitalizados requerem uma conversão para serem inseridos nesse contexto. Ambos os tipos de objetos digitais trazem desafios específicos em termos de preservação, já que os objetos digitalizados podem sofrer perda de qualidade durante o processo de conversão, demandam tempo e recursos para serem adequadamente processados, e exigem atenção para garantir que metadados e direitos autorais sejam devidamente gerenciados (CHAPMAN, 2004; BREEDING, 2014). Por outro lado, os nato-digitais dependem de *softwares* e formatos que podem se tornar obsoletos com o tempo.

---

[7] A digitalização é o processo de conversão de materiais físicos em digitais (SOCIETY OF AMERICAN ARCHIVISTS, 2024), como a digitalização de prontuários médicos em papel ou exames radiológicos em filme, que são convertidos em formatos digitais específicos.

2.2 Dados e Metadados

Os objetos digitais são construídos sobre dados, conforme definido pela Escola Nacional de Administração Pública (ENAP), "[os dados] representam fatos através de um conjunto de caracteres primitivos e isolados, geralmente representados através de textos, números, imagens, sons ou vídeos". SAYÃO; SALES (2012) definem que a ideia de dado inclui tanto os objetos digitais simples, compostos por um único arquivo, identificador e metadados, quanto os objetos digitais complexos, que são a combinação de diversos objetos digitais formando uma unidade discreta. Um exemplo típico é uma página web, que agrupa diferentes elementos digitais. Nesse contexto, as bases de dados são definidas como coleções estruturadas de registros ou dados armazenados em sistemas computacionais, de modo a facilitar a organização e recuperação de informações.

Dados podem ser categorizados em estruturados, semiestruturados e não estruturados. Dados estruturados são organizados em um formato padronizado, como tabelas em um banco de dados, onde elementos como nomes de pacientes ou resultados de exames são facilmente processados. Dados semiestruturados, como arquivos *Extensible Markup Language* (XML) ou *JavaScript Object Notation* (JSON), possuem uma organização flexível que permite variações em sua estrutura, comum em prontuários médicos eletrônicos. Dados não estruturados, como anotações médicas, transcrições de consultas e imagens de exames, não seguem um formato definido e são mais complexos de organizar e analisar. À medida que as organizações dependem cada vez mais desses dados, o seu valor como ativo pode ser estabelecido com mais clareza (DAMA, 2017).

Uma peça fundamental para gerenciar dados como ativos é o uso de metadados confiáveis. Embora sejam geralmente definidos como "dados sobre os dados", a distinção entre o que constitui dado e metadado pode variar dependendo do contexto. Ou seja, informações que podem ser consideradas como dados em uma situação podem ser vistas como metadados em outra. (FHIR [...], 2020) Os metadados representam os significados dos dados.

GRÁCIO (2012) amplia essa compreensão, descrevendo os metadados como "um conjunto de dados, chamados de elementos, cujo número é variável, de acordo com um padrão adotado, que descreve o recurso, possibilitando a um usuário ou a um

mecanismo de busca acessar e recuperar esse recurso". Ainda destaca que a utilização de padrões de metadados "permite a troca de informações entre instituições que utilizam o mesmo padrão ou até mesmo entre aquelas que utilizam padrões diferentes". Para a *National Information Standards Organization* (NISO), metadados são "informação estruturada que descrevem, explicam, localizam, ou ainda possibilitam que um recurso informacional seja fácil de recuperar, usar ou gerenciar" e são criados para diferentes finalidades.

Os metadados descritivos são criados para descrever e facilitar a descoberta de conteúdo digital, fornecendo informações sobre o que o conteúdo trata e ajudando na sua localização. Já os metadados administrativos são utilizados de forma abrangente para se referir às informações utilizadas para gerenciar o conteúdo ou relacionadas com a sua criação. Podem se referir aos metadados técnicos, que estão embutidos nos arquivos digitais e descrevem suas características técnicas; aos metadados de preservação, cujo objetivo é auxiliar no gerenciamento à longo prazo e nas técnicas futuras de migração ou emulação dos arquivos digitais; e aos metadados de direito, que detalham os direitos de propriedade intelectual associados ao conteúdo. Há ainda os metadados estruturais, criados para descrever os relacionamentos de partes dos recursos entre si (RILEY, 2017). A Tabela 2.1 resume os diferentes tipos de metadados.

Tabela 2.1 – Tipos de metadados segundo a NISO

Tipo		**Função**	**Usos primários**	**Exemplos de elementos**
Metadados descritivos		Para encontrar ou entender um recurso	Descoberta; Exibição; Interoperabilidade	Título; autor; assunto; gênero; data de publicação
Metadados administrativos	**Metadados técnicos**	Para descodificar e compilar arquivos	Interoperabilidade; Gerenciamento de objetos digitais; Preservação	Tipo de arquivo; Tamanho do

				arquivo; Criação data/hora; Esquema de compressão
	Metadados de preservação	Gerenciamento de arquivos a longo prazo;	Interoperabilidade; Gestão de objetos digitais; Preservação	Soma de Verificação; Código h*ash*
	Metadados de direito	Direitos de propriedade intelectual associados ao conteúdo	Interoperabilidade; Gestão de objetos digitais	Direitos autorais; Licenças; Titular dos direitos
Metadados estruturais		Relacionamentos de partes de recursos entre si	Navegação	Localização sequencial na hierarquia; Páginas em sequência; Índice com apontadores para o início de seções

Fonte: CASTRO; SANTOS (2018), RILEY (2017)

Dessa forma, concluímos que os padrões de metadados estabelecem uma forma comum de estruturar e entender dados, e inclui princípios e regras de implementação para utilizá-los (UNIVERSITY OF PITTSBURGH LIBRARIES, 2023). Por exemplo, o *Dublin Core* (DC) é um dos padrões de metadados mais amplamente utilizados e reconhecidos, oferecendo um conjunto de 15 elementos descritivos para recursos digitais. Nas ciências, temos padrões como *Darwin Core* (DwC) e *Access to Biological Collection Data* (ABCD) para dados biológicos, e o *Astronomy Visualization Metadata Standard* (AVM) para imagens astronômicas. Nas ciências

sociais, a *Data Documentation Initiative* (DDI) é utilizada para dados observacionais, enquanto nas artes e humanidades, o *Text Encoding Initiative Guidelines* (TEI) é um padrão para representação de textos digitais.

Como padrão de metadados de preservação digital, o *PREservation Metadata: Implementation Strategies* (PREMIS), fornece um Dicionário de Dados composto por unidades semânticas, com o objetivo de definir os elementos mais comuns para a maioria dos repositórios de preservação (PREMIS EDITORIAL COMMITTEE, 2015). Como define CAPLAN (2009), "uma unidade semântica é um fragmento de informação", que descreve características essenciais para os sistemas conhecerem e exportarem. O Modelos de Dados PREMIS definir quatro entidades: objeto, eventos, direitos e agente.

A entidade *objeto* representa os elementos que realmente são armazenados e gerenciados em um repositório de preservação, e podem ser categorizados em arquivo, *bitstream*, representação e entidade intelectual. A entidade *evento* contém informações sobre as ações que afetam os *objetos* ao longo do seu ciclo de vida no repositório. Os *direitos* documentam as informações sobre direitos e permissões referentes aos objetos, para que os repositórios realizem as ações necessárias para sua preservação. Por fim os *agentes* inclui os atores (pessoas, organizações, softwares ou hardwares) que têm funções nos *eventos*, nas declarações de *direitos* e nos *objetos ambiente* (CAPLAN, 2009). Cada padrão deve ser projetado para atender às necessidades de descrição e organização de recursos em seu domínio específico.

No que diz respeito à gestão de sistemas de saúde, padrões como o *Health Level 7* (HL7) Internacional, por exemplo, tem o objetivo de garantir a interoperabilidade entre diferentes sistemas e facilitar a troca, integração, compartilhamento e recuperação de informações de saúde eletrônicas. A interoperabilidade pode ser dividida em dois níveis principais: sintática e semântica. A interoperabilidade sintática assegura que os dados possam ser transmitidos e lidos por diferentes sistemas de forma consistente, utilizando formatos padronizados, como os definidos pelo HL7, por exemplo.

Já a interoperabilidade semântica vai além, garantindo que o significado e o propósito dos dados trocados sejam compreendidos da mesma maneira por todos os sistemas envolvidos (MELLO; MESQUITA; VIEIRA, 2015). Alguns exemplos desse tipo são: o

padrão *Open Eletronic Health Records* (openEHR), que permite que os profissionais da área médica definam arquétipos para o gerenciamento da informação clínica, incluindo sua hierarquia, agregação e terminologias, de forma colaborativa em uma plataforma aberta de gerenciamento do conhecimento clínico (HEARD, 2018); e a terminologia *Systematized Nomenclature of Medicine - Clinical Terms* (SNOMED-CT), que fornece uma linguagem comum que permite que sistemas diferentes compreendam os mesmos conceitos médicos, como procedimentos, estruturas corporais, organismos, sintomas, substâncias, diagnósticos, tratamentos, etc. e sem ambiguidades (SHAHPORI; DOIG, 2010). Ambos foram definidos pela Portaria nº 2.703, de 31 de agosto de 2011, e são adotados no Brasil.

O padrão HL7, com sua arquitetura de documentos clínicos conhecida como *Clinical Document Architecture* (CDA), é utilizado para especificar a estrutura e a semântica de documentos clínicos, garantindo que sejam legíveis tanto por máquinas quanto por humanos. Isso garante que a troca de informações, como solicitações e resultados de exames, seja realizada de maneira uniforme entre os sistemas, preservando a integridade e o significado dos dados (SALES; BENTES PINTO, 2019).

De acordo com a HIMSS (2006), a interoperabilidade, nessa definição mais ampla, abrange a movimentação consistente dos dados entre sistemas, mantendo o significado dos dados intacto, além de assegurar a apresentação uniforme das informações, os controles de segurança e a confidencialidade dos pacientes. Isso permite que os prestadores de saúde realizem uma análise integrada da situação do paciente, incluindo seus riscos, sintomas e contexto, o que melhora o processo e aumenta a precisão na prevenção, no tratamento e no diagnóstico da doença ou condição de saúde (D'AGOSTINO et al., 2020, tradução nossa).

Outros padrões de informação foram definidos pela Portaria nº 2.703, como o padrão *Logical Observation Identifiers Names and Codes* (LOINC) para a codificação de exames laboratoriais; a norma ISBT 128 para codificação de dados de identificação das etiquetas de produtos relativos ao sangue humano, de células, tecidos e produtos de órgãos; a especificação de integração *Patient Identifier Cross-Referencing* (IHE-PIX) para o cruzamento de identificadores de pacientes de diferentes sistemas de informação; e outras classificações que serão utilizadas para suporte à interoperabilidade dos sistemas de saúde: Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde (CID),

Classificação Internacional de Atenção Primária (CIAP-2) e Classificação brasileira hierarquizada de procedimentos médicos (CBHPM).

No que diz respeito à descrição de imagens médicas, o *Digital Imaging and Communications in Medicine* (DICOM) é amplamente utilizado, e permite o compartilhamento de imagens de diagnóstico entre sistemas que utilizam o padrão, como tomografia computadorizada, ressonância magnética, ultrassonografia, mamografia, dentre outros (BOTÃO, 2019). Esse padrão engloba diversos aspectos de imagiologia médica[8] digital, desde a codificação dos dados que compõem uma imagem, parâmetros de visualização destas imagens, formato de arquivamento em disco, serviços para a comunicação de imagens e informações através de redes de computadores (SALES; BENTES PINTO, 2019).

O padrão Troca de Informações na Saúde Suplementar (TISS) foi estabelecido como um padrão obrigatório no Brasil para trocas eletrônicos de dados de atenção à saúde de beneficiários de planos. O objetivo é padronizar as ações administrativas, subsidiar as ações de avaliação e acompanhamento econômico, financeiro e assistencial das operadoras de planos privados de assistência à saúde e compor o Registro Eletrônico de Saúde (ANS, 2024). Além disso, para fins de padronização e unificação da terminologia de definição de serviços de prestadores de saúde suplementar, a ANS definiu a Terminologia Unificada de Saúde Suplementar (TUSS), que deve ser usada obrigatoriamente em todas as transações suportadas pelo padrão TISS (LUIS et.al, 2020).

À nível global, o padrão de metadados ISBT (*International Society of Blood Transfusion*) 128 foi criado para estabelecer uma uniformização mundial de rotulagem de hemocomponentes, tecidos e produtos de terapia celular que, por meio de código de barras, permite, entre outros benefícios, uma identificação ampla das características dos produtos e do seu local de coleta. O ISBT 128 especifica: um sistema de numeração de doação, que assegura a identificação global e exclusiva; a

---

[8] Refere-se ao conjunto das técnicas e dos procedimentos que permitem obter imagens do corpo humano com fins clínicos ou científicos. A radiologia, a termografia médica, a endoscopia, a microscopia e a fotografia médica fazem parte destas técnicas.

informação a ser transferida, por meio de tabelas de referência internacionalmente acordadas; um banco de dados internacional de referência do produto; as estruturas de dados nas quais essa informação é colocada; um sistema de codificação em barras para a transferência da informação no rótulo do produto; *layouts* tipo padrão para o rótulo do produto; um padrão de referência para uso em mensagens eletrônicas (MINISTÉRIO DA SAÚDE, 2012).

A adoção de padrões facilita não só o compartilhamento de informações entre os sistemas e sua interoperabilidade, mas também se enquadra como uma estratégia estrutural que busca facilitar a execução das outras estratégias de preservação e maximizar sua eficiência (FORMENTON; GRACIOSO; CASTRO, 2015). Graells (2020) complementa que um dos pontos importantes da preservação digital é a utilização de padrões de metadados apropriados e formatos de arquivos específicos.

Nesse sentido, o gerenciamento bem executado de metadados e sua adesão a padrões reconhecidos promovem um entendimento consistente dos recursos de dados e um desenvolvimento interorganizacional mais eficaz. As organizações obtêm mais valor de seus ativos de dados se os dados forem de alta qualidade, portanto dependem e são críticos para a governança. (BARBOSA; SHAYER LYRA, 2021)

2.3 Governança de Dados

O *Data Governance Institute* (DGI) define a governança de dados como "um sistema de direitos de decisão e responsabilidades para processos relacionados a informações, executados de acordo com modelos acordados que descrevem quem pode tomar quais ações, com quais informações, quando, sob que circunstâncias e usando quais métodos." O DAMA *International* (2017) pontua ainda que a governança de dados "fornece os princípios, políticas, processos, estrutura, métricas e supervisão necessários para gerenciar os dados como um ativo e orientar as atividades de gerenciamento de dados em todos os níveis", isso representa uma separação inerente de responsabilidades entre supervisão e execução (Figura 2.1).

Figura 2.1 – Relação Governança de Dados e Gestão de Dados

Fonte: DAMA (2017), traduzido pela autora

De modo complementar, a Gestão de Dados engloba as operações diárias de programas e de organizações no contexto de estratégias, políticas, processos e procedimentos que tenham sido estabelecidos pelo órgão. Seu foco é na garantia da eficácia, ou seja, cumprir as ações prioritárias, e na eficiência das operações, buscando realizar as atividades da maneira mais otimizada possível em termos de custo-benefício. (BARBOSA; SHAYER LYRA, 2021)

Nesse contexto, o *Data Management Body of Knowledge* (DMBOK) atua como um guia abrangente que oferece uma estrutura detalhada para a prática da gestão de dados. Ele aborda diversos aspectos, e serve como um recurso essencial para profissionais de dados, fornecendo orientações sobre os processos, políticas, métricas e estruturas necessárias para uma gestão eficaz dos dados como um ativo organizacional. A Roda do DAMA-DMBOK2 (segunda edição) (Figura 2.2) entra como uma representação visual das Áreas de Conhecimento em Gestão de Dados. Ela coloca a Governança de Dados no centro e equilibra outras 10 áreas, são elas:

1. **Arquitetura de Dados**: Define o plano para gerenciar os ativos de dados, e estabelece requisitos de dados estratégicos e projetos para atender a esses requisitos.
2. **Modelagem e *Design* de Dados**: Processo de descoberta, análise, representação e comunicação dos requisitos de dados.

3. **Armazenamento e Operações de Dados**: Inclui o *design*, implementação e suporte de dados armazenados para maximizar seu valor. As operações oferecem suporte durante todo o ciclo de vida dos dados, desde o planejamento até a disposição dos dados.
4. **Segurança de Dados**: Garante que a privacidade e a confidencialidade dos dados sejam mantidas, que os dados não sejam violados e que os dados sejam acessados de forma apropriada.
5. **Integração e Interoperabilidade de Dados**: Inclui processos relacionados ao movimento e consolidação de dados dentro e entre armazenamentos de dados, aplicativos e organizações.
6. **Gerenciamento de Documentos e Conteúdo**: Inclui atividades de planejamento, implementação e controle usadas para gerenciar o ciclo de vida de dados e informações encontradas em uma variedade de mídias não estruturadas, especialmente documentos necessários para atender aos requisitos legais e regulatórios de conformidade.
7. **Dados de Referência e Mestre**: Inclui a reconciliação contínua e a manutenção de dados críticos compartilhados essenciais para permitir o uso consistente em sistemas da versão mais precisa, oportuna e relevante da verdade sobre entidades de negócios essenciais.
8. ***Data Warehousing* e *Business Intelligence***: Inclui processos de planejamento, implementação e controle para gerenciar dados de suporte à tomada de decisões e permitir que os trabalhadores do conhecimento obtenham valor dos dados por meio de análises e relatórios.
9. **Metadados**: Inclui atividades de planejamento, implementação e controle para permitir acesso a metadados integrados de alta qualidade, incluindo definições, modelos, fluxos de dados e outras informações críticas para entender dados e os sistemas por meio dos quais são criados, mantidos e acessados.
10. **Qualidade de Dados**: Inclui o planejamento e implementação de técnicas de gerenciamento de qualidade para medir, avaliar e melhorar a adequação dos dados para uso dentro de uma organização.

Figura 2.2 – Roda do DAMA-DMBOK2

Fonte: DAMA (2017), traduzido pela autora

Ainda que o foco principal do DAMA não seja a preservação digital, suas áreas do conhecimento, conforme definidas pelo DMBOK, oferecem uma base para práticas de gestão de dados que podem ser complementares a iniciativas de preservação digital. Considerando essa relação, Silva (2023) realizou uma análise para correlacionar cada área de conhecimento do DMBOK com os conceitos de preservação e curadoria digital, utilizando como referência o modelo *Open Archival Information System* (OAIS) e o Modelo de Ciclo de Vida de Curadoria do *Digital Curation Centre* (DCC). Esta análise será explorada com mais detalhes no tópico 2.5.1.

No contexto da saúde, uma governança de dados eficiente deve fornecer ferramentas, definir conceitos e estruturar ações de modo a permitir que instituições de saúde gerem dados de maior qualidade, confiabilidade e segurança. Além disso, deve adotar um modelo de informação em sistemas de saúde, que seja sustentável a longo prazo e que assegure a privacidade e a confidencialidade dos dados pessoais dos cidadãos

(D'AGOSTINO et al., 2017, tradução nossa). A seguir, serão discutidas as principais iniciativas do governo brasileiro relacionadas à gestão e governança em saúde digital.

2.4 Saúde Digital

A Organização Mundial de Saúde (OMS) define a saúde digital como "o campo de conhecimento e prática associado ao desenvolvimento e uso de tecnologias digitais para melhorar a saúde. A saúde digital inclui os consumidores digitais, dispositivos inteligentes e equipamentos conectados. Também abrange outros usos de tecnologias digitais para saúde, como a Internet das Coisas, inteligência artificial, *big data* e robótica." (WHO, 2021, tradução nossa). Países como Canadá, Austrália, Estados Unidos, Inglaterra, Escócia, e Suécia vêm investindo sistematicamente em infraestrutura, sistemas, serviços, recursos humanos e em modelos de organização para tornar a e-Saúde parte do cotidiano da saúde e uma estratégia de sua melhoria (SWEDEN, 2010; AUSTRALIA, 2011; ENGLAND, 2011; OFFICE OF THE NATIONAL COORDINATOR FOR HEALTH INFORMATION TECHNOLOGY, 2011; SCOTLAND, 2011; PHILIPPINES, 2012; S. SOUTH AFRICA, 2012; CANADA HEALTH INFOWAY, 2016; *apud* BRASIL, 2017).

A estratégia digital global da OMS, conforme definido em WHO (2021) enfatiza que os dados de saúde devem ser classificados como dados pessoais sensíveis e que exigem um padrão de segurança elevado. Portanto, ela destaca a necessidade de uma base legal e regulatória sólida para proteger a privacidade, confidencialidade, integridade e disponibilidade dos dados e o processamento de dados pessoais de saúde, e para lidar com segurança cibernética, construção de confiança, responsabilidade e governança, ética, equidade, capacitação e alfabetização, garantindo que dados de boa qualidade sejam coletados e posteriormente compartilhados para apoiar o planejamento, comissionamento e transformação de serviços.

Existem no mercado diferentes soluções tecnológicas, plataformas e *softwares*, cada um deles com funcionalidades específicas, como prontuários eletrônicos, captação, armazenamento e compartilhamento de imagens de exames, emissão de prescrição eletrônica, assinaturas digitais e outras tantas. Entretanto, justamente em um

universo de aplicações desconexas, a falta de governança e interoperabilidade se torna um risco (LIN et al., 2023).

Nesse contexto de crescente informatização da saúde, os Registros Eletrônicos em Saúde (RES) representam uma coleção abrangente de informações de saúde de um paciente que podem ser acessadas e gerenciadas por várias instituições e profissionais de saúde ao longo do tempo. Para capturar, armazenar, apresentar, transmitir ou imprimir esses registros é utilizado o Sistema de Registro Eletrônico de Saúde (S-RES) (ISO, 2005), que possibilita ainda a interoperabilidade entre diferentes sistemas e regiões. Um componente crucial desse sistema é o Prontuário Eletrônico do Paciente (PEP), que se refere ao conjunto de informações clínicas detalhadas de um paciente em uma única instituição de saúde.

Com o objetivo de avaliar e atestar aspectos de qualidade, segurança e privacidade de um S-RES, a Sociedade Brasileira de Informática em Saúde (SBIS), em parceria com o CFM, desenvolveu o processo de certificação de sistemas informatizados em saúde, instituído pela Resolução CFM nº 1821/2007. A Certificação de S-RES SBIS é composta por três estágios que representam o nível de maturidade do *software*, e possui uma série de requisitos de estrutura, conteúdo, funcionalidades e segurança (KIATAKE et al., 2020).

Entre esses requisitos estão o controle de versão do *software* para rastreabilidade e investigações sobre alterações, assim como resgate dos códigos-fonte correspondentes; a identificação e autenticação de usuários, estabelecendo métodos de autenticação (como senhas com controle mínimo de segurança, biometria e certificados digitais) para controlar o acesso individual aos dados de saúde conforme os papeis aos quais o usuário possui, assim como dos métodos de armazenamento desses dados ou parâmetros utilizados no processo de autenticação; *backups* completos (*full*) e o armazenamento de metadados associados aos dados contendo informações suficientes para que possam ser restauradas com integridade; o sistema deve gerar registros de auditoria de forma ininterrupta, garantindo que todas as operações no sistema sejam documentadas, incluindo criação, consulta, acréscimo ou substituição de registros de saúde, importação e exportação de dados, tentativas de autenticação e a realização de *backups* (MARQUES et al., 2016).

O movimento de informatização dos Sistemas de Informação em Saúde (SIS) no Brasil, iniciado na década de 1970, marcou o começo da centralização de dados de

saúde pelo Governo Federal, que estabeleceu padrões nacionais para a coleta e armazenamento de informações, visando enfrentar a inconsistência nos dados nos níveis estadual e municipal (MACHADO; TAVARES, 2023). Esse processo evoluiu com a criação do Departamento de Informática do SUS (DATASUS) em 1991, que, desde então, fornece suporte em sistemas de informação e informática para o planejamento e gestão dos serviços de saúde (BRASIL, 2023, MEIRELLES; CUNHA, 2020).

Sob responsabilidade do DATASUS, o governo brasileiro lançou a Estratégia de Saúde Digital para o Brasil para 2028 (ESD28), que pretende consolidar e avançar as políticas de saúde digital no país. A ESD28 sistematiza e consolida o trabalho realizado nessas áreas ao longo da última década, materializado em diversos documentos. As publicações relevantes para a saúde digital no contexto brasileiro estão representadas na Figura 2.3.

Figura 2.3 – Contextualização da saúde digital no Brasil

Fonte: Adaptado de BRASIL (2020a)

No Brasil, o Sistema de Saúde Suplementar e o SUS formam a base do sistema de saúde brasileiro. O Sistema de Saúde Suplementar no Brasil é regulada pelo poder público através da Agência Nacional de Saúde (ANS) e refere-se ao conjunto de serviços de saúde oferecidos por meio de planos de saúde privados, que são contratados de forma voluntária pelos cidadãos ou por empresas que oferecem esses planos como benefício para seus funcionários (PIETROBON; PRADO; CAETANO, 2008). Este sistema complementa o SUS, o sistema público de saúde do Brasil,

criado pela Constituição Federal de 1988, com o objetivo de garantir acesso universal, integral e gratuito aos serviços de saúde para toda a população brasileira.

O SUS está estruturado em diferentes níveis de atenção e assistência à saúde, que variam conforme a complexidade dos serviços prestados: atenção primária, atenção secundária e terciária. Como descrito em FRASÃO; RIBEIRO (2022), na Atenção Primária à Saúde (APS) são realizados ações e atendimentos voltados à prevenção e promoção à saúde, como consultas, exames simples e acompanhamento de doenças crônicas. A atenção especializada, por sua vez, é dividida em dois elementos: atenção secundária e terciária, que são de, respectivamente, média e alta complexidade (ambulatorial e especializada hospitalar).

A média complexidade é composta por serviços especializados encontrados em hospitais e ambulatórios e envolve atendimento direcionado para áreas como pediatria, ortopedia, cardiologia, oncologia, neurologia, psiquiatria, ginecologia, oftalmologia entre outras especialidades. Já os serviços de alta complexidade incluem procedimentos de maior custo, que demandam tecnologia mais avançada, como tratamentos oncológicos, cirurgias cardiovasculares, transplantes e atendimentos em unidades de terapia intensiva (UTI), além de partos de alto risco e cirurgias de grande porte. (FRASÃO; RIBEIRO, 2022; MINISTÉRIO DA SAÚDE, 2010)

Em acordo com essa estrutura, a Política Nacional de Informação e Informática em Saúde (PNIIS), atualizada pela última vez através da Portaria GM/MS nº 1.768/2021, define os princípios e diretrizes que orientam tanto o setor público quanto o privado na integração dos sistemas de informação em saúde, na promoção da transparência, segurança e acesso às informações de saúde pela população. A Gestão Arquivística de Documentos (GAD), os RES, os sistemas de arquivos e os repositórios institucionais são explicitamente citados nas diretrizes da PNIIS. (MEIRELLES; CUNHA, 2020)

Além disso, a política atribui ao Ministério da Saúde (MS), juntamente com as secretarias de saúde estaduais, municipais e do Distrito Federal, a responsabilidade de promover uma Política de Governança de Dados em Saúde em conformidade com a Lei Geral de Proteção de Dados (LGPD), além de desenvolver e gerir padrões e estratégias de saúde digital. Nesse contexto, no que diz respeito ao tratamento de dados na área da saúde, incluindo dados digitais, a LGPD determina no artigo 11, parágrafo 4º que:

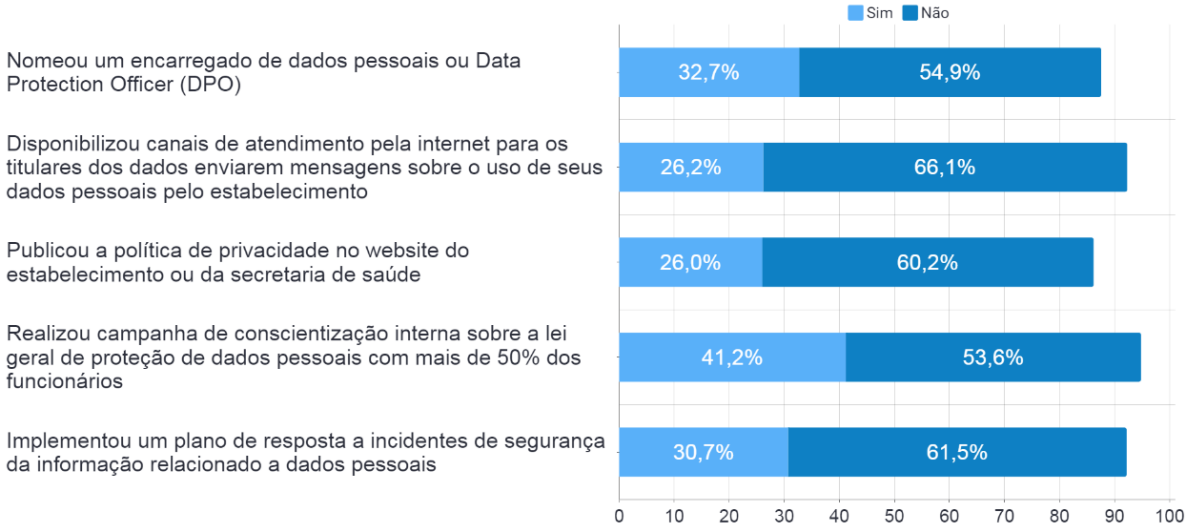
> É vedada a comunicação ou o uso compartilhado entre controladores de dados pessoais sensíveis referentes à saúde com objetivo de obter vantagem econômica, exceto nas hipóteses relativas a prestação de serviços de saúde, de assistência farmacêutica e de assistência à saúde, desde que o tratamento de dados de saúde não seja para a prática de seleção de riscos na contratação de qualquer modalidade, assim como na contratação e exclusão de beneficiários, incluídos os serviços auxiliares de diagnose e terapia, em benefício dos interesses dos titulares de dados, e para permitir: I - a portabilidade de dados quando solicitada pelo titular; ou II - as transações financeiras e administrativas resultantes do uso e da prestação dos serviços de que trata este parágrafo. (adaptado de BRASIL, 2018a)

A LGPD obriga ainda a nomeação pública de um encarregado de dados pessoais, ou *Data Protection Officer* (DPO), em organizações públicas e privadas. O papel principal do DPO é garantir que a organização processe os dados pessoais de seus funcionários, clientes, fornecedores ou qualquer outra pessoa (também chamada de titulares de dados[9]) em conformidade com as regras de proteção de dados aplicáveis (EDPS, 2004). No entanto, uma pesquisa de 2022 do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (CETIC), cuja missão é monitorar a adoção das tecnologias de informação e comunicação (TIC) no Brasil, apontou que somente 32,7% dos estabelecimentos de saúde nomeou um encarregado de dados.

---

[9] O titular de dados é a pessoa natural a quem se referem os dados pessoais que são objeto de tratamento (BRASIL, 2018a)

Figura 2.4 – Estabelecimentos de saúde, por medidas adotadas em relação à LGPD



Fonte: CETIC (2022)

Como mostrado na Figura 2.4, a pesquisa também indicou falhas na adoção de boas práticas e governança de dados nos estabelecimentos. O artigo 9º da legislação assegura o direito do titular ao acesso facilitado às informações sobre o tratamento de seus dados, enfatizando que estas devem ser disponibilizadas de maneira clara, adequada e ostensiva. No entanto, apenas 29,2% dos estabelecimentos ofereceram canais de atendimento para acesso fácil dos titulares aos seus dados, e apenas 29% publicaram suas políticas de privacidade.

Quanto à conscientização interna sobre a LGPD, o artigo 46 estipula a adoção de medidas de segurança técnicas e administrativas para proteger os dados pessoais, o que inclui a capacitação dos funcionários sobre as melhores práticas de proteção de dados. O artigo 47, por sua vez, destaca a importância de avaliações periódicas dos procedimentos internos relacionados ao tratamento de dados pessoais, exigindo a revisão e atualização das políticas de segurança da informação e programas de conscientização. Na prática, apenas 41,2% dos estabelecimentos realizaram campanhas de conscientização que alcançasse mais de 50% dos funcionários.

Esses dados ressaltam a necessidade de melhora nas práticas de governança de dados e da conscientização interna sobre a proteção de dados pessoais nos estabelecimentos de saúde. É essencial atender aos requisitos da LGPD para assegurar a segurança e privacidade dos dados dos titulares. Nesse sentido, o Ministério da Saúde desenvolveu ferramentas-chave no programa Conecte SUS,

como a Rede Nacional de Dados em Saúde (RNDS), o e-SUS Atenção Primária (e-SUS APS) e a ESD28, como parte dos esforços para implementar a PNIIS. (MINISTÉRIO DA SAÚDE, 2021)

Na perspectiva de um estabelecimento de saúde, a RNDS oferece serviços de interoperabilidade em saúde no território brasileiro, instituída pela Portaria GM/MS nº 1.434, de 28 de maio de 2020, cuja integração de informações será feita de forma gradativa até a concretização da Rede como via única de interoperabilidade nacional em saúde. A proposta do RNDS relaciona-se à interoperabilidade sintática por meio da construção de um repositório único de informações em saúde por meio da troca de informações em todos os níveis de atenção à saúde, vigilância e gestão em saúde. (MACHADO; TAVARES, 2023) No Brasil, a RNDS é o meio pelo qual um estabelecimento em saúde disponibiliza informação que será consumido por outro, seguindo o padrão *Fast Healthcare Interoperability Resources* (FHIR) ("Introdução | RNDS", [s.d.]).

Já o e-SUS APS atua como uma estratégia para reestruturar e qualificar a gestão das informações da APS em todo o Brasil. Os objetivos principais dessa iniciativa são individualizar e integrar os registros de saúde dos cidadãos, reduzir o retrabalho na coleta de dados, qualificar as informações em saúde, e otimizar a gestão e a coordenação do cuidado. Para isso, conta com dois grandes sistemas para armazenamento e coleta de dados: o Sistema de Informação em Saúde para a Atenção Básica (SISAB), responsável pelo processamento e disseminação de informações, e o Sistema e-SUS APS, composto por *softwares* e aplicativos que auxiliam na coleta de dados clínicos e administrativos. (BRASIL, 2023)

A estratégia inclui o Prontuário Eletrônico do Cidadão (PEC), um *software* que potencialmente oferece maior consistência nas informações, se preenchidas adequadamente, mas pode resultar em obsolescência dos sistemas e suas funcionalidades. Essa estratégia está relacionada com outra abordagem para a obtenção de dados em saúde em formato integrado e interoperável, diferente da abordagem adotada pela RNDS, que por sua vez tentar replicar a consistência de dados que o uso de um único sistema potencialmente oferece, o que implica em custos de adaptação e desenvolvimento contínuos. O Brasil adota uma combinação dessas duas estratégias. (MACHADO; TAVARES, 2023)

Outro trabalho relevante para construção da ESD28 é a Estratégia e-Saúde para o Brasil, de 2017. Esta estratégia realiza uma análise e comparação dos recursos existentes no país relacionados à aspectos como governança, recursos organizacionais, sistemas e serviços de saúde, padrões e interoperabilidade, infraestrutura e recursos humanos. Com base nessa análise, são recomendadas uma série de ações estratégicas para preencher as lacunas identificadas. As ações estratégicas e os objetivos mapeados estão resumidos na Tabela 2.2.

Tabela 2.2 – Resumo das Ações Estratégicas

Ação Estratégica	**Objetivo**
Reduzir a fragmentação das iniciativas no SUS e aprimorar a governança da estratégia de e-Saúde.	• Fortalecer as instâncias de governança da informação no SUS/MS e promover o alinhamento das suas ações. • Organizar o ambiente de implantação da Visão de e-Saúde. • Construir um arcabouço institucional capaz de orquestrar as inúmeras ações em curso no âmbito da e-Saúde.
Fortalecer a intersetorialidade de governança de e-Saúde	• Dar uma dimensão nacional à Visão de e-Saúde para o Brasil, integrando programas e agregando recursos de todos os setores de governo, da sociedade civil e da iniciativa privada para viabilizar a Visão de e-Saúde proposta neste documento.
Elaborar o marco legal de e-Saúde no País	• Adequar o marco legal para suportar as inovações do campo da atenção à saúde
Definir e implantar uma arquitetura para a e-Saúde	• Construir uma arquitetura de e-Saúde composta ao menos pelos seguintes blocos reutilizáveis: Modelos de informação e artefatos de conhecimento, modelo para interoperabilidade de serviços de terminologia, cadastros nacionais de identificação, consentimento, serviços e sistemas de segurança e privacidade, e arquitetura de intercâmbio de informações de saúde.

Definir e implantar os sistemas e serviços de e-Saúde	• Garantir que a infraestrutura computacional, necessária para a implantação da Visão de e-Saúde para o Brasil, esteja disponível e evolua de acordo com as necessidades e oportunidades advindas da evolução tecnológica.
Criar arquitetura de referência para sustentação dos serviços de infraestrutura	• Desenvolver e estabelecer uma arquitetura de referência para infraestrutura, visando à plena sustentação dos serviços de TIC para e-Saúde.
Criar a certificação em e-Saúde para trabalhadores do SUS	• Implantar um processo de certificação baseado na formação e atualização profissional em e-Saúde no SUS, como fator estruturante e orientador da qualificação dos profissionais do SUS para a implantação da Visão de e-Saúde.
Promover a facilitação do acesso à informação em saúde para a população	• Implantar recursos de e-Saúde que promovam a facilitação do acesso da população à informação em saúde de qualidade, estimulando o acesso à informação em saúde para a população.

Fonte: Adaptado de BRASIL (2017)

Essas ações foram reafirmadas, atualizadas e expandidas no documento da ESD28, que descreve o conjunto de atividades a serem executadas e os recursos necessários para a implementação da Visão de Saúde Digital. O Plano de Ação baseia-se nas ações propostas no Plano de Ação, Monitoramento e Avaliação de Saúde Digital (PAM&A) para o Brasil 2019-2023, publicado em 2020, principalmente aquelas estabelecidas no Programa Conecte SUS. Das sete prioridades do Plano de Ação da ESD28, cinco se destacam como as mais relevantes para esta pesquisa, são elas: 1. Governança e Liderança para a ESD; 2. Informatização dos três níveis de atenção; 5. Formação e Capacitação de Recursos Humanos; 6. Ambiente de Interconectividade; e 7. Ecossistema de Inovação.

As ações de Governança e Liderança para a ESD incluem a consolidação da governança, segurança, privacidade e confidencialidade de dados alinhados com os preceitos da LGPD, infraestrutura robusta, financiamento sustentável e estável e colaboração multissetorial com definição de normas claras e públicas. A estratégia estabelece bases sólidas para o gerenciamento, acesso e preservação de dados de saúde a longo prazo, alinhando-se com princípios essenciais dessas áreas e cria um

ambiente propício para iniciativas de curadoria e preservação digital em saúde, mesmo que não as aborde diretamente.

A informatização de estabelecimentos de saúde do país, prioridade 2 do ESD, abrange ações como a viabilização do acesso à internet para os estabelecimentos assistenciais de saúde em todo o território nacional; e a informatização (a adoção de sistemas de informação) de todas as unidades de saúde, equipes de Estratégia de Saúde da Família[10] (ESF), equipes de APS do país e de todos os estabelecimentos de Atenção Especializada e Hospitalar do território nacional (BRASIL, 2020a). Além dos benefícios esperados pela estratégia, que incluem melhoria do atendimento mediante acesso às informações de saúde, maior segurança nos dados e fortalecimento da continuidade do cuidado e potencialização da capacidade de ação do Governo de formulação de políticas públicas, também implicam na melhoria das práticas de curadoria e preservação de dados.

Essas ações permitem a centralização dos dados de saúde em sistemas integrados, como prontuários eletrônicos e bases de dados unificadas. Isso facilita a padronização dos dados, garantindo que eles sejam coletados e armazenados de maneira consistente e conforme normas e padrões internacionais. A informatização também melhora o acesso e a recuperação de dados, essencial para a preservação digital, ao permitir a localização e o acesso rápido e seguro às informações. A segurança e integridade dos dados são fortalecidas por mecanismos de segurança incorporados nos sistemas de informação, que protegem contra acessos não autorizados e perda de dados. É possível ainda implementar estratégias de preservação digital, objeto desta pesquisa e que busca garantir a acessibilidade dos dados ao longo do tempo. Por fim, a interoperabilidade é promovida por essas ações, permitindo que os dados sejam compartilhados entre diferentes estabelecimentos de saúde e plataformas.

---

[10] A ESF emprega equipes multidisciplinares, que incluem médicos, enfermeiros, e agentes comunitários, para prestar cuidados de saúde, desde a promoção da saúde até a reabilitação, focando no cuidado integrado e direcionado à população em áreas geográficas específicas ("Estratégia Saúde da Família", 2015).

A Prioridade 5 da ESD, que foca na formação e capacitação de recursos humanos para a Saúde Digital, relaciona-se diretamente com as atividades de preservação digital. A capacitação de profissionais é essencial para garantir que os serviços e aplicativos de Saúde Digital sejam utilizados de forma plena, não apenas na prática clínica, mas também na gestão e na preservação dos dados de saúde. Os profissionais da saúde podem assumir a responsabilidade pela curadoria de dados clínicos, desde que possuam competências para tal. A criação da categoria profissional do Informata em Saúde e o reconhecimento formal dessa área do conhecimento, ações previstas na ESD, contribuem para o desenvolvimento e aplicação de estratégias de preservação digital. Isso inclui a implementação de melhores modelos de representação de dados, padrões de segurança, e práticas de preservação que assegurem a continuidade do cuidado e a integridade dos dados de saúde.

A criação de um ambiente de interconectividade em saúde enfatiza a importância de ações relacionadas à interoperabilidade de dados e padronização de sistemas e informações em saúde, o que favorece a criação de um ambiente para futuras iniciativas de preservação digital, uma vez que formatos e padrões abertos facilitam a migração e o acesso aos dados a longo prazo. Outra ação relevante diz respeito ao desenvolvimento do Repositório de Terminologias de Saúde (RTS): manter um repositório de terminologias padronizadas ajuda a preservar o contexto e o significado das informações ao longo do tempo, de forma que os dados permaneçam compreensíveis e utilizáveis no futuro, mesmo quando as tecnologias e práticas médicas evoluírem.

Por fim, o ecossistema de inovação, estimula a criação e expansão dos serviços integrados da RNDS e do Lago de Dados[11] de Saúde e o desenvolvimento de iniciativas em IoT e *Big Data*, que implica na geração e compartilhamento de um grande volume de dados de saúde. Esse ecossistema pretende promover um ambiente

---

[11] Conforme definido na ESD28, "o Lago de Dados (do inglês, *Data lake*) consiste em uma arquitetura tecnológica capaz de armazenar e disponibilizar um alto volume de dados sem necessidade de tratamento prévio, em altíssima velocidade, permitindo que haja um repositório centralizado para compartilhamento de informações com ferramentas de acesso e análise em tempo real" (BRASIL, 2020a).

colaborativo e interconectado onde o SUS, as organizações de saúde privadas, empresas de tecnologia, centros de pesquisa, universidades e outros atores possam compartilhar dados e experiências, além de testar e avaliar novos modelos, padrões e tecnologias (BRASIL, 2020a). Para que esse ambiente de inovação seja sustentável e capaz de suportar a evolução constante da tecnologia e das necessidades de saúde, as ações de preservação digital tornam-se também fundamentais.

2.5 Curadoria Digital

Os significados de "curadoria" variam de acordo com o contexto histórico e o campo de atuação. No entanto, em sua essência, a curadoria está associada ao cuidado, à preservação e à seleção de objetos, sejam eles obras de arte, patrimônio cultural, coleções científicas ou mesmo a proteção dos interesses de pessoas incapazes de tomar decisões. A curadoria é essencial em contextos organizacionais específicos, como bibliotecas, arquivos, museus e instituições similares, onde a gestão e a preservação de coleções são de grande importância.

Na área da Ciência da Computação, o termo Curadoria Digital foi utilizado pela primeira vez no evento *"Digital Curation: digital archives, libraries and e-Science seminar"*[12], em 2001, e já apontava para a necessidade de se fomentar discussões e determinar atividades de ação conjunta para a gestão de acervos digitais, no intuito de "compartilhar experiências práticas de curadoria digital no setor de bibliotecas digitais, arquivos e ciências eletrônicas" (DIGITAL PRESERVATION COALITION *apud* BRAYNER 2018, p. 54).

O evento lançou também a Digital Preservation Coalition (DPC), uma organização do Reino Unido, cujo objetivo é criar uma comunidade global inclusiva focada na preservação sustentável de ativos digitais, capacitando seus membros com boas práticas de criação, gerenciamento e preservação de materiais digitais e fornecendo uma variedade de ferramentas práticas para auxiliar nesse processo. Desde então, a

[12] Curadoria Digital: seminário de Arquivos Digitais, Bibliotecas e e-Ciência", em tradução livre.

elaboração de conceitos teóricos, criação de planos estratégicos e aplicação de modelos operacionais no campo da curadoria digital vêm se estabelecendo de forma cada vez mais ampla e consistente para o gerenciamento, acesso e utilização dos acervos e dados eletrônicos. (BRAYNER, 2018)

O DCC se destaca também como um ator relevante nesse cenário, desempenhando um papel fundamental no apoio à comunidade de curadoria digital e preservação. Lançado em 2004, concentra seus esforços na resolução de desafios relacionados à curadoria digital e preservação de longo prazo, oferecendo serviços de suporte compartilhados a instituições de ensino superior do Reino Unido e expandindo sua atuação globalmente. Com foco na gestão de dados de pesquisa e na promoção de práticas que visam tornar os dados acháveis, acessíveis, interoperáveis e reutilizáveis (FAIR)[13], o DCC desempenha um papel vital no desenvolvimento de estratégias e modelos operacionais para diversas áreas de pesquisa e educação. (DCC, 2004)

O DCC define a curadoria digital como "o gerenciamento e preservação de dados/informações digitais a longo prazo, e envolve a manutenção, preservação e agregação de valor aos dados digitais ao longo de seu ciclo de vida". Sayão e Sales (2012) complementam essa ideia enfatizando que o principal desafio recai na necessidade de se preservar não somente o conjunto de dados, mas de preservar, sobretudo, a capacidade que ele possui de transmitir conhecimento para uso futuro das comunidades interessadas.

Nas próximas seções, serão apresentados e analisados alguns dos principais modelos de ciclos de vida que são relevantes para a curadoria digital. Esses modelos são representações gráficas ou conceituais que ilustram as fases pelas quais os dados passam desde a sua criação até a sua preservação ou descarte.

---

[13] Os Princípios FAIR (Findable, Accessible, Interoperable, Reusable) são atributos desenvolvidos e validados pela comunidade científica que possibilitam usar e citar os dados corretamente.

2.5.1 Digital Curation Lifecycle Model – DCC

O Modelo de Ciclo de Vida de Curadoria do DCC[14] (Figura 2.5) é uma contribuição relevante que complementa uma série de padrões que fornecem *frameworks* para a gestão de objetos digitais. A adoção de uma abordagem baseada no ciclo de vida fornece uma visão de alto nível que pode ser usada em conjunto com modelos de referência, estruturas e padrões para auxiliar no planejamento de atividades em níveis mais detalhados, além de assegurar que todos os estágios necessários sejam devidamente identificados e planejados, permitindo a implementação das ações necessárias na sequência apropriada. (HIGGINS, 2008)

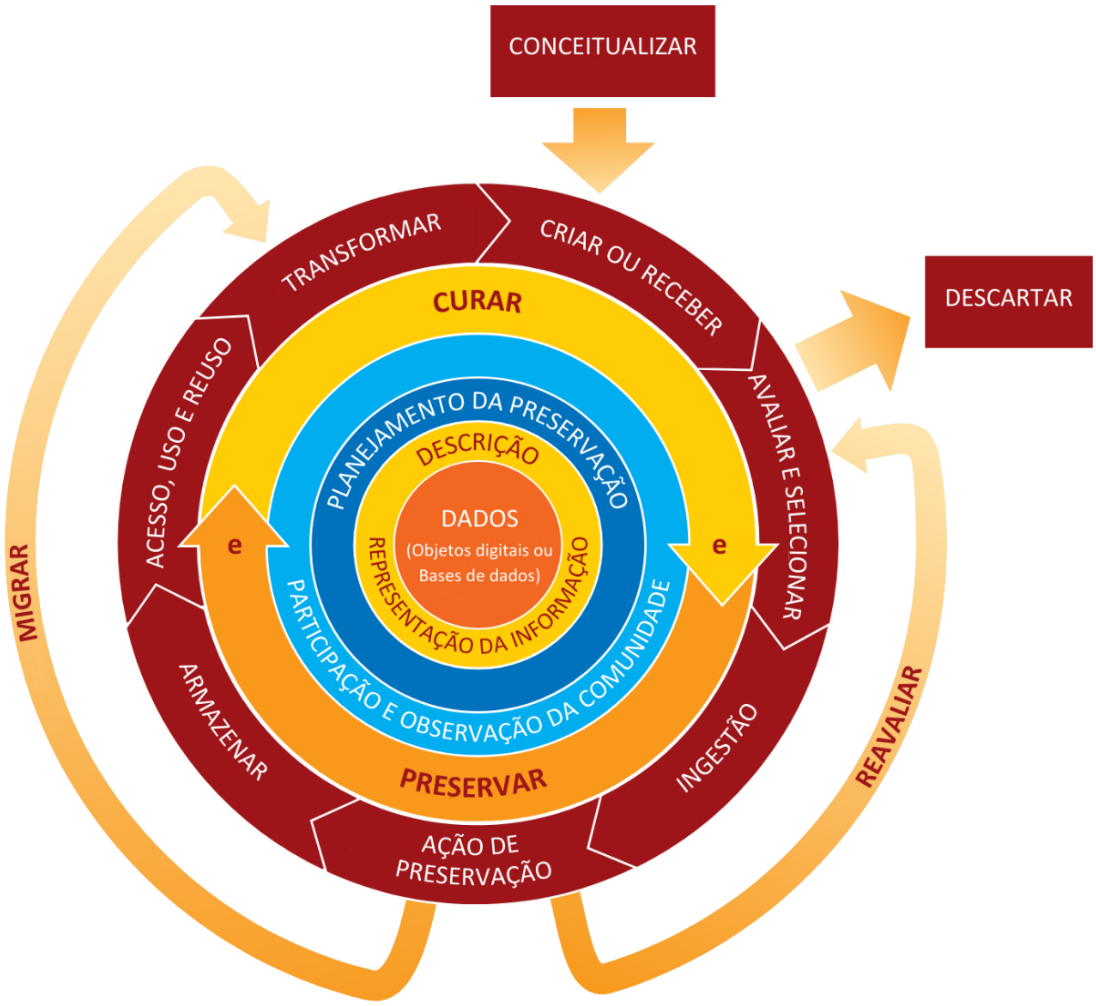
De acordo com Higgins (2008), o modelo de Ciclo de Vida de Curadoria do DCC oferece uma visão abrangente das etapas essenciais para a curadoria e preservação de dados, desde a concepção inicial até a entrega final. A autora destaca a importância desse modelo para o planejamento e organização das atividades de uma organização, garantindo que todas as etapas necessárias ocorram na ordem correta. Ele também auxilia na identificação de funções específicas, na definição de responsabilidades e na criação de padrões e tecnologias para a implementação. Além disso, ajuda a identificar etapas adicionais necessárias e ações não aplicáveis em determinados contextos, contribuindo para a documentação adequada de processos e políticas.

O ciclo de vida da curadoria coloca os dados no centro de suas operações. Isso inclui tanto objetos digitais simples, que são itens digitais discretos, como arquivos de texto, imagens ou arquivos de áudio, juntamente com seus identificadores relacionados e metadados; quanto objetos digitais complexos, que são objetos digitais discretos criados pela combinação de vários outros objetos digitais, como sites. Além disso, o modelo também abrange bases de dados, definidas como coleções estruturadas de registros ou dados armazenados em sistemas computacionais.

---

[14] The DCC Curation Lifecycle Model, em inglês.

Figura 2.5 – O modelo de Ciclo de Vida da Curadoria Digital do DCC



Fonte: HIGGINS (2008), traduzido pela autora

Uma série de ações englobam todo o ciclo de vida no modelo do objeto digital, guiando o processo de curadoria digital. Para transmitir essa ideia de presença contínua, essas ações estão representadas graficamente como anéis concêntricos envolvendo os objetos de dados (SAYÃO; SALES, 2012). Essas ações incluem a **Descrição e Representação da Informação**, que atribui metadados administrativos, descritivos, técnicos, estruturais e de preservação, utilizando padrões apropriados, para garantir uma descrição adequada e controle a longo prazo, bem como a coleta e atribuição de informações de representação necessárias para compreender e renderizar tanto o material digital quanto os metadados associados.

Além disso, o **Planejamento de Preservação** enfatiza a necessidade de planejar a preservação ao longo do ciclo de vida da curadoria do material digital, o que envolve

planos para gestão e administração de todas as ações do ciclo de vida da curadoria. A **Observação e Participação da Comunidade** envolvem a vigilância sobre as atividades da comunidade apropriada e participar do desenvolvimento de padrões compartilhados, ferramentas e *software* adequado. Por fim, estar ciente e realizar ações de gestão e administração planejadas para promover a curadoria e preservação ao longo do ciclo de vida da curadoria são ações de **Curar e Preservar**.

O modelo também define as ações sequenciais, que devem ser realizadas continuamente todo o tempo que o dado estiver sob curadoria. A ação de **Conceitualização** concebe e planeja a criação do dado, incluindo os métodos de captura e opções de armazenamento. A **Criação ou Recebimento** diz respeito à criação do dado, incluindo metadados administrativos, descritivos, estruturais e técnicos. A preservação de metadados também pode ser adicionada na etapa de criação. Essa ação também inclui a recepção do dado de acordo com políticas documentadas de coleta, recebidas pelos criadores do dado, outros arquivos, repositórios ou centro de dados, e, se necessário, a atribuição de metadados apropriados.

Também entre as ações que devem continuar ciclicamente, estão **Avaliar e Selecionar**, em que é feita a avaliação dos dados e seleção para curadoria e preservação a longo prazo; **Ingestão** (ou Arquivamento), para transferir os dados para um arquivo, repositório, centro de dados ou outro custodiante. Ambos os processos devem seguir orientações documentadas, políticas ou requisitos legais.

A **Ação de Preservação** consiste na realização de ações para garantir a preservação a longo prazo e a manutenção da sua credibilidade. As ações de preservação devem garantir que os dados permaneçam autênticos, confiáveis e utilizáveis, mantendo sua integridade. As ações incluem limpeza de dados, validação, atribuição de metadados de preservação, atribuição de informações de representação e garantia de estruturas de dados ou formatos de arquivo aceitáveis. O **Armazenamento** envolve armazenar os dados de forma segura, seguindo padrões relevantes. As ações de **Acesso, Uso e Reuso** pretendem garantir que os dados sejam acessíveis tanto para os usuários designados quanto para os reutilizadores, diariamente. Isso pode ser feito na forma de informações publicamente disponíveis. Podem ser aplicáveis controles de acesso e procedimentos de autenticação. Finalmente, a **Transformação** compreende a criação

de novos dados a partir do original, seja por migração para um formato diferente ou pela criação de subconjuntos, por seleção ou consulta, gerando novos resultados que podem ser publicados.

O modelo do DCC também estabelece algumas ações ocasionais, são atividades que podem ocorrer em diferentes momentos do ciclo de vida dos dados, conforme surgem necessidades específicas. Essas ações incluem o **Descarte**, para descartar dados que não foram selecionados para curadoria e preservação a longo prazo, conforme as políticas documentadas, orientações ou requisitos legais; a **Reavaliação**, para reavaliar os dados que falham nos procedimentos de validação para uma nova avaliação e resseleção; e a **Migração**, para migrar os dados para um formato diferente, pode ser feito para adequar-se ao ambiente de armazenamento ou para garantir a imunidade dos dados à obsolescência de *hardware* ou *software*.

As ações previstas no modelo de Ciclo de Vida de Curadoria do DCC complementam e enriquecem a abordagem mais ampla do DMBOK, que por sua vez fornece uma estrutura que aborda várias áreas de conhecimento da gestão de dados. Ambos os modelos adotam uma visão abrangente do ciclo de vida dos dados, mas o fazem de maneiras distintas e complementares. O DCC, com seu modelo cíclico específico para curadoria digital, pode ser visto como um aprofundamento de alguns aspectos abordados de forma mais geral no DMBOK.

A área de Armazenamento e Operação de dados do DMBOK tem como objetivo a disponibilidade dos dados, desde a criação até o descarte, garantindo a integridade e o desempenho dos dados (DAMA, 2017). Esse foco se alinha com as ações do modelo DCC que visam preservar e garantir o acesso contínuo e seguro aos dados ao longo de seu ciclo de vida. O DMBOK destaca a importância de escolher o melhor modelo de armazenamento com base nos tipos de dados, algo que ressoa com a etapa de Preservação do DCC, onde se busca garantir que o dado original, autêntico e íntegro seja armazenado de maneira segura e acessível. Além disso, a preocupação do DMBOK com a migração dos dados para novos ambientes tecnológicos encontra um paralelo direto na ação de Migração do DCC, que visa exatamente essa movimentação e adaptação dos dados conforme as mudanças tecnológicas (SILVA, 2023).

Em termos de segurança, o DMBOK foca na proteção dos dados contra acessos não autorizados e criptografia para proteger informações sensíveis, enquanto o DCC

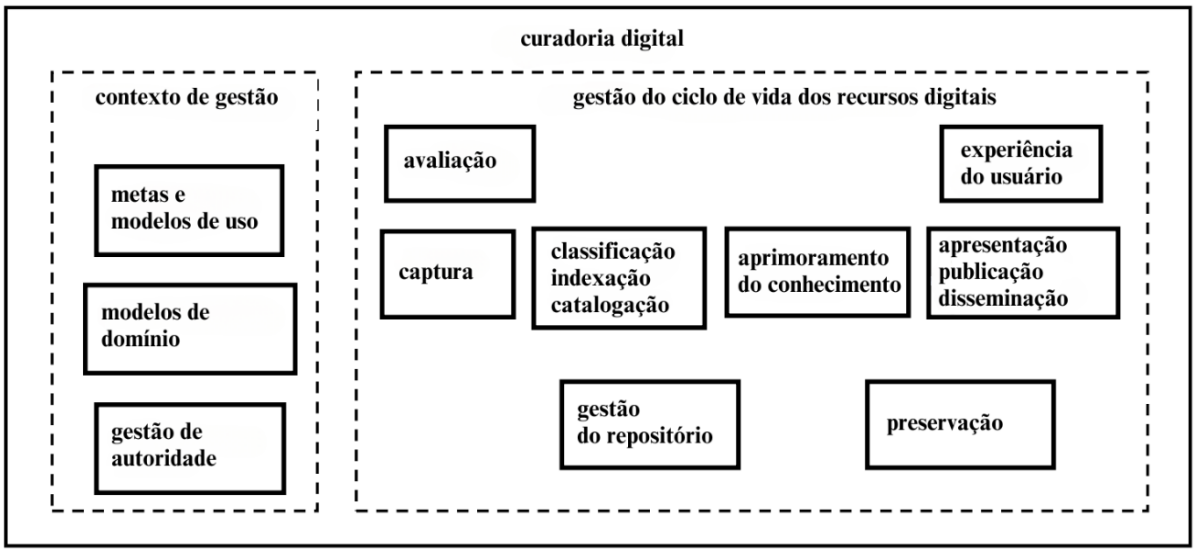
incorpora aspectos de segurança em várias ações para garantir a integridade dos dados. Como destaca Silva (2013), a ação de Acesso, Uso e Reuso traz ainda recomendações para a aplicação de controle robusto e procedimento de autenticação. Os Metadados são reconhecidos como cruciais tanto no DMBOK quanto no DCC, com este último enfatizando a Descrição e Representação da Informação, que envolve a atribuição de metadados em várias categorias. Com a integração dessas duas abordagens é possível construir uma estrutura sólida para a gestão e curadoria de dados, garantindo que as necessidades específicas de preservação sejam atendidas ao mesmo tempo em que se mantém a conformidade com as melhores práticas de gestão de dados.

2.5.2 Digital Curation Unit Model e Extended Digital Curation Lifecycle Model (DCC&U) – DCU

O modelo estendido de ciclo de vida de curadoria é uma proposta que visa expandir e aprimorar o modelo de curadoria digital proposto pelo DCC, incorporando elementos-chave da abordagem orientada a processos do *Digital Curation Unit* (DCU)[15]. Essa abordagem compreende os processos envolvidos na curadoria digital, como descrito por Constantopoulos e Dallas (2007) e ilustrado na Figura 2.6.

[15] Disponível em: http://www.dcu.gr/en/

Figura 2.6 – Modelo do DCU



Fonte: CONSTANTOPOULOS; DALLAS (2007), traduzido pela autora

Os processos definidos na gestão do ciclo de vida dos recursos digitais dependem de três processos de suporte, que capturam o contexto da curadoria digital e produzem recursos que podem ser vistos como ativos digitais curados (CONSTANTOPOULOS; DALLAS, 2007). São eles:

- **Metas e modelos de uso** (*goal and usage models*): Capturam tanto as intenções do criador (metas), quanto os padrões de uso dos recursos de uma determinada classe pelos interagentes (modelos de uso).
- **Modelos de domínio** (*domain models*): Produz ou refina representações de conhecimento sobre o domínio de interesse.
- **Gestão de autoridade** (*authority management*): Lida com o controle de vocabulários (termos geográficos, períodos históricos, moléculas químicas, espécies biológicas etc.) usados por convenção para denotar conceitos, propriedades e relações. Essa é uma fase contextual que merece atenção por parte dos curadores, porque a gestão de autoridade evolui ao longo do tempo e, assim, se faz relevante manter atualizada a representação o objeto digital por meio de uma descrição que traga o conceito e suas relações, otimizando assim, o contexto e o domínio do conhecimento atrelado ao objeto digital (SILVA; SIEBRA, 2017)

Do ponto de vista do ciclo de vida da informação, a curadoria digital abrange uma série de processos voltados para alcançar (a) a confiabilidade dos recursos digitais, (b) organização, arquivamento e preservação de longo prazo, e (c) serviços de valor agregado e novos usos para os recursos. (CONSTANTOPOULOS; DALLAS, 2007) Esses processos incluem:

- **Avaliação** (*appraisal*): É o processo de desenvolver critérios (determina quais fatores são importantes ao avaliar recursos potenciais, inclui considerações como utilidade, segurança da informação, qualidade do conteúdo e precisão da informação) e selecionar recursos (identifica os recursos que têm potencial de uso e reuso e que são valiosos para a finalidade pretendida) que podem se tornar parte de processos subsequentes de curadoria.

- **Captura** (*ingesting*): A ingestão de recursos digitais é o processo de integrá-los ao processo de curadoria. Isso pode ser feito através da criação de gravações digitais de imagens, sons, textos e dados; da conversão de gravações analógicas em mídia física para formatos digitais; ou da importação de conteúdo digital existente de outras fontes. Para a organização que coleta e gerencia esses recursos, os materiais ingeridos se tornam sua principal fonte de informação.

- **Classificação, indexação e catalogação** (*classification, indexing and cataloguing*): São processos fundamentais para a organização e acesso a recursos digitais. Produzem os índices lógicos, temáticos e relacionados aos usos potenciais do material. É importante ressaltar que os próprios índices podem ser considerados recursos digitais e podem ser criados durante o processo ou importados de outras fontes.

- **Aprimoramento do conhecimento** (*knowledge enhancement*): É o processo de enriquecer o entendimento sobre as entidades, situações e eventos do mundo real representados por recursos digitais, seu contexto e domínio. Isso inclui anotar documentos com entidades relevantes, descrever formalmente situações e eventos, e vincular documentos entre si de acordo com seu conteúdo. Similar à indexação, o aprimoramento

pode resultar na criação de novos recursos digitais que são autônomos e complementares aos recursos originais.

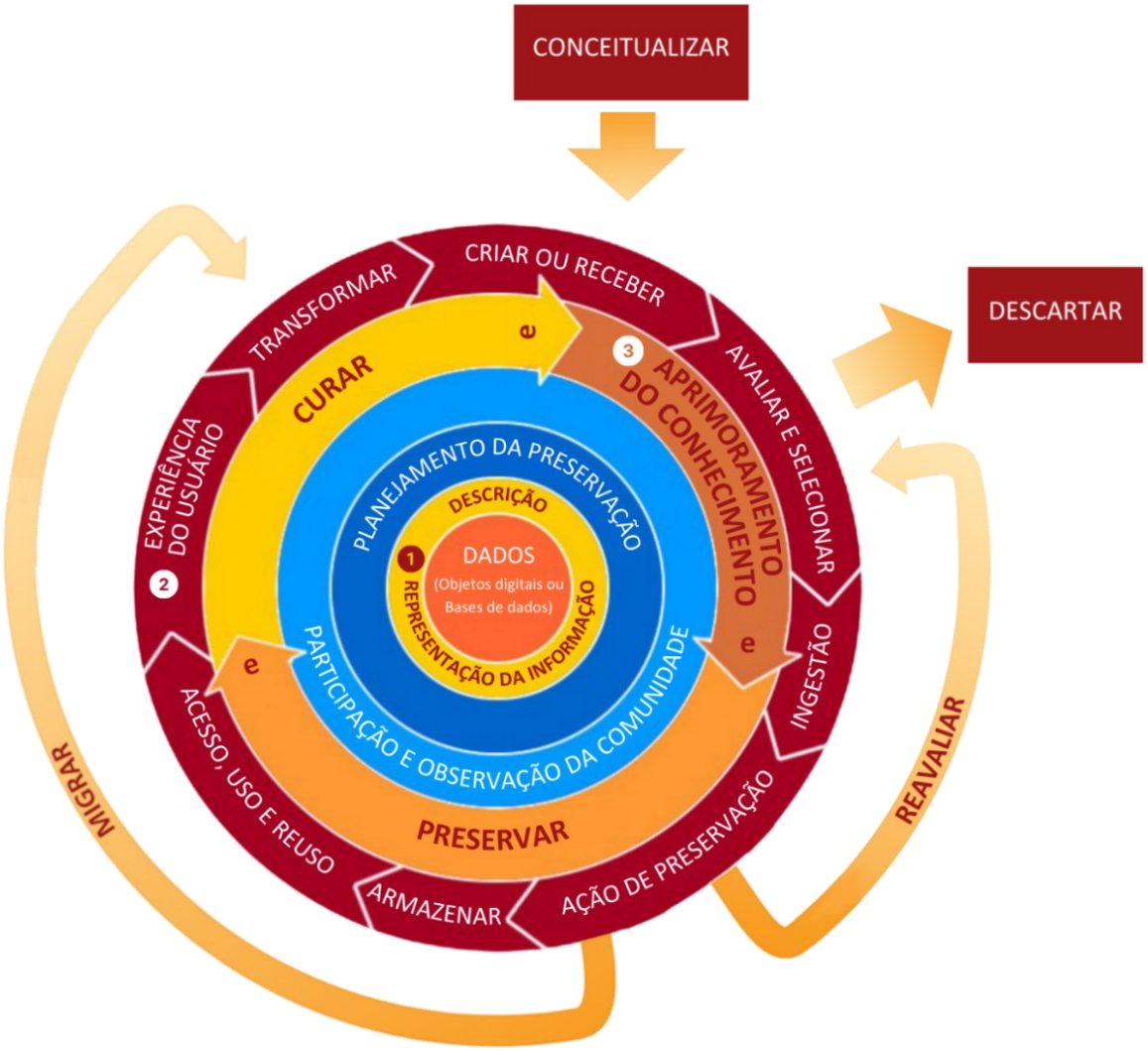
- **Apresentação, publicação e disseminação** (*presentation, publication and dissemination*): São processos que envolvem a geração de novos artefatos (científicos, acadêmicos, artísticos etc.) a partir de recursos digitais primários ou secundários existentes.
- **Experiência do usuário** (*user experience*): Este processo captura a interação entre usuários e recursos, assim como os efeitos dessa interação.
- **Gerenciamento do repositório** (*repository management*): Todos os recursos digitais são armazenados, organizados e gerenciados em um repositório. Esse processo está relacionado tanto com repositórios físicos (centralizados ou distribuídos) quanto com repositórios virtuais, assim como com seus mecanismos de acesso.
- **Preservação** (*preservation*): Processo destinado a proteger contra riscos à longevidade, recorrentes de causas físicas (incidentes com os meios de armazenamento e desastres ambientais) ou da evolução tecnológica. (CONSTANTOPOULOS; DALLAS (2007); CONSTANTOPOULOS, DALLAS et al. (2009); SILVA; SIEBRA (2017))

O modelo do DCU destaca algumas atividades que não foram consideradas no modelo original do DCC. As adições podem ser vistas na Figura 2.7 e incluem a adição de 1) a inclusão de vocabulários controlados, como nomes geográficos, períodos históricos, moléculas químicas, espécies biológicas, entre outros, que são usados por convenção para denotar conceitos, propriedades e relações (ação para todo ciclo de vida); 2) o registro da experiência do usuário ao acessar os dados (ação sequencial); e 3) conhecimento aos repositórios de recursos digitais, o que representa uma nova maneira de interpretar ou combinar os recursos primários com conhecimentos pré-existentes (ação para todo ciclo de vida) (CONSTANTOPOULOS; DALLAS *et al.*, 2009).

Os autores do modelo consideram estes aspectos como sendo cruciais para o processo de curadoria digital, visto que proporcionam uma visão mais completa e detalhada do ciclo de vida da curadoria digital, considerando as necessidades e

interações dos usuários e as características específicas dos dados objetos digitais, além de estimular uma discussão mais ampla sobre nessa área.

Figura 2.7 – DCC&U: Modelo estendido do Ciclo de Vida da Curadoria Digital



Fonte: Adaptado de CONSTANTOPOULOS; DALLAS (2009), traduzido pela autora

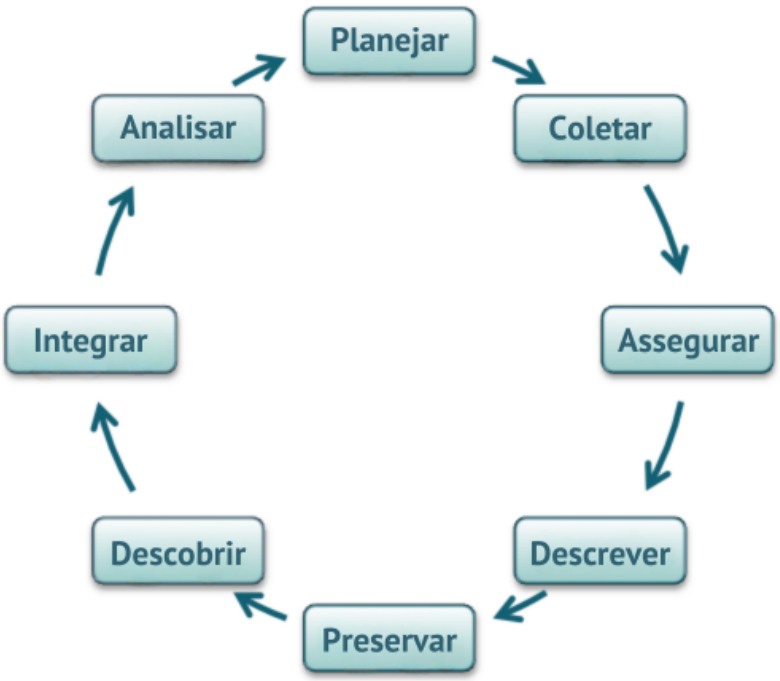
2.5.3 Data Lifecycle – DataONE

Inspirado pela importância da preservação e uso responsável dos dados científicos, o projeto *Data Observation Network for Earth* (DataONE) se destaca como um pioneiro na promoção das melhores práticas em gestão de dados. Fundamentado na premissa da relevância dos dados para o avanço da pesquisa, o modelo do DataONE

tem sido amplamente adotado como uma estrutura fundamental para a curadoria de dados em diversas disciplinas (ARAÚJO et al., 2019).

Este modelo, com suas oito etapas delineadas (Figura 2.8), é reconhecido pela sua flexibilidade, permitindo que projetos adaptem o ciclo de vida da curadoria de acordo com suas necessidades específicas. Como observado por STRASSER et al. (2012), embora o modelo apresente uma sequência linear, é importante ressaltar que a realidade da curadoria de dados pode ser mais fluida, com projetos frequentemente revisitando etapas ou seguindo caminhos não lineares para alcançar seus objetivos.

Figura 2.8 – O ciclo de vida dos dados da DataONE



Fonte: PLALE; KOUPER (2017), traduzido pela autora

- **Planejar** (*Plan*): É crucial elaborar um plano de gestão de dados que seja revisado regularmente e adaptado conforme necessário ao longo do projeto. Nessa etapa, deve ser considerado a definição dos dados a serem coletados e analisados, a escolha de um repositório de dados apropriado, a organização e gestão eficiente dos dados, a descrição detalhada deles, o compartilhamento com colegas e a comunidade científica, a implementação de medidas de preservação de curto prazo e a consideração dos recursos financeiros e institucionais disponíveis.

- **Coletar** (*Collect*): Nessa etapa os dados são transformados em formato digital. É importante considerar os métodos e documentação antes da coleta de forma a garantir sua usabilidade no futuro. Nessa etapa é recomendado criar um modelo para uso durante a coleta de dados, descrever o conteúdo dos arquivos de dados, organizar de forma consistente os dados em um arquivo, usar o mesmo formato ao longo do arquivo, utilizar caracteres simples para nomes de variáveis, nomes de arquivos e dados; preferir *software* e *hardware* não proprietários, atribuir nomes descritivos aos arquivos, manter os dados brutos originais e criar tabelas de parâmetros e de locais.

- **Assegurar** (*Assure*): Diz respeito à garantia e controle de qualidade durante todas as etapas do processo de dados, incluindo coleta, entrada e análise. Isso pode ser garantido ao descrever as condições de coleta que possam afetar a qualidade dos dados, utilizar indicadores de qualidade para verificar valores estimados e inseridos manualmente. Manter a consistência no formato dos dados e identificar valores discrepantes por meio de análises estatísticas; comunicar a qualidade dos dados usando códigos ou metadados e identificar e corrigir valores ausentes; utilizar conjuntos de dados semelhantes para identificar problemas potenciais e estar atento a problemas adicionais durante a análise e interpretação dos dados, também são atividades que compreendem essa etapa.

- **Descrever** (*Describe*): A documentação abrangente de dados (*i.e.* metadados) é fundamental para a compreensão futura dos dados. Para que os dados possam ser facilmente descobertos, compreendidos ou utilizados efetivamente é recomendado uma descrição detalhada do contexto do arquivo de dados, do contexto científico, das informações sobre os parâmetros utilizados, da qualidade dos dados, da equipe e das partes interessadas. Esses metadados devem ser gerados em um formato de metadados comumente usados pela comunidade científica mais relevante.

- **Preservar** (*Preserve*): Nessa etapa é crucial trabalhar com um repositório de dados especializado para orientação sobre a preparação de

metadados, a seleção de formatos de arquivo adequados e a garantia de serviços adicionais para futuros usuários. Deve ser considerado utilizar terminologia padrão para identificar os dados; considerar as políticas legais e de privacidade, com as permissões necessárias e adequadamente licenciados; e documentar informações de proveniência aos dados, indicando responsáveis, contexto do projeto e histórico de revisões.

- **Descobrir** (*Discover*): Dados potencialmente uteis são localizados e obtidos, em conjunto com informações relevantes sobre o dado (metadados).
- **Integrar** (*Integrate*): Dados de fontes distintas são combinados para formar um conjunto homogêneo de dados que pode ser facilmente analisado.
- **Analisar** (*Analyze*): Os dados são analisados.

A descoberta, integração, análise e visualização de dados são apoiadas por uma variedade de ferramentas, que auxiliam os pesquisadores. Quando novos conjuntos de dados são criados a partir de conjuntos de dados existentes ou de seus elementos individuais, é importante documentar e rastrear essa relação.

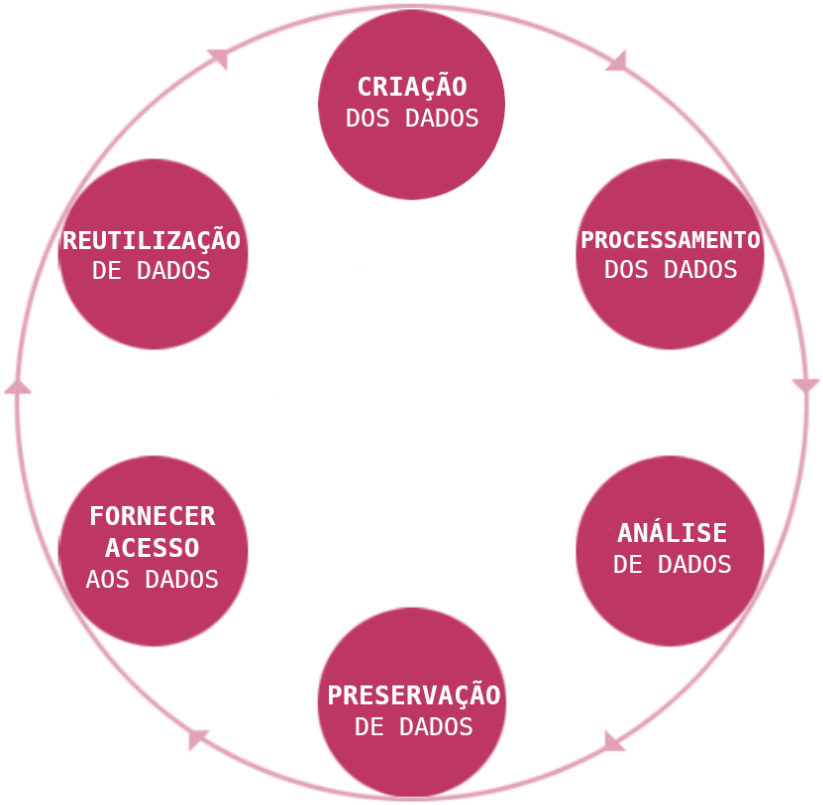
2.5.4 UK Data Archive Data Lifecycle

O *UK Data Archive*[16] é um centro nacional de *expertise* em arquivamento de dados no Reino Unido. Ele fornece um modelo de ciclo de vida de dados como um auxílio para pesquisadores que estão considerando como a gestão de dados se relaciona com o ciclo de vida de um projeto de pesquisa (BALL, 2012). O modelo define seis etapas que podem ser vistas na Figura 2.9.

[16] Disponível em: https://www.data-archive.ac.uk/

A **Criação dos dados** envolve o planejamento do gerenciamento de dados, incluindo formatos e armazenamento, bem como o planejamento do consentimento para compartilhamento. Além disso, nesta fase, busca-se também por dados existentes, coletam-se novos dados por meio de experimentação, observação, medição ou simulação, e são capturados e criados metadados para fornecer informações contextuais sobre os dados. A etapa seguinte é o **Processamento de dados**. Isso envolve a entrada de dados, digitalização, transcrição e tradução, bem como a verificação, validação e limpeza dos dados para garantir sua qualidade. Quando necessário, os dados são anonimizados para proteger a privacidade dos participantes.

Figura 2.9 – O ciclo de vida dos dados da UK Data Archive



Fonte: EYNDEN (2013), traduzido pela autora.

Na etapa de **Análise de dados** os dados coletados são preparados para análise e interpretação. As atividades relacionadas com essa etapa englobam a interpretação dos dados, a derivação de novos dados a partir das análises realizadas, a produção de resultados de pesquisa e a autoria de publicações científicas com base nos resultados obtidos. Em **Preservação de Dados**, o foco está na garantia da longevidade e acessibilidade dos dados ao longo do tempo. Isso envolve a migração dos dados para

o melhor formato disponível, backups regulares e armazenamento seguro dos dados, a criação de metadados e documentação de preservação detalhada e o trabalho contínuo de preservação e curadoria dos dados para garantir sua integridade e utilidade futura.

Em **Fornecer acesso aos dados**, o objetivo é disponibilizar os dados para uso por outros pesquisadores, profissionais e o público em geral. Isso envolve distribuir e compartilhar os dados, controlar o acesso para garantir a segurança e a conformidade com as políticas de uso, estabelecer direitos autorais, quando aplicável, e promover ativamente o uso dos dados para incentivar a colaboração e a inovação na comunidade acadêmica e além dela.

Por fim, a **Reutilização de Dados** tem como foco incluir a realização de pesquisas de acompanhamento, análises secundárias dos dados, revisões de pesquisa, examinar minuciosamente os resultados e conclusões obtidos, e utilizar os dados para fins educacionais. Essa etapa promove o aproveitamento máximo dos dados coletados, disponibilizando os dados para serem utilizados por outros pesquisadores e educadores (BALL, 2012).

2.5.5 Digital Content Lifecycle Model – DigitalNZ

A *DigitalNZ*[17] é um serviço administrado pela Biblioteca Nacional da Nova Zelândia e financiado pelo governo da Nova Zelândia que hospeda mídia digital relacionada ao país desde 2008. As organizações parceiras incluem bibliotecas, museus, galerias, departamentos governamentais, meios de comunicação social e grupos comunitários. O conteúdo inclui fotografias, mapas, vídeos, obras de arte, reportagens e gravações

[17]Disponível em: https://digitalnz.org/

de áudio. Os metadados são estruturados e disponibilizados por meio de uma *Application Programming Interface* (API)[18] de uso gratuito. (DIGITALNZ, 2006b)

Em 2009, baseado no Ciclo de Vida de Conteúdo Digital desenvolvido pela equipe da DigitalNZ, foi lançado o *Make it digital*, que busca auxiliar e explicar por meio de dicas práticas quaisquer aspectos para tornar o material digital disponível e útil. O ciclo de vida abaixo (Figura 2.10) enfatiza isso por meio de sete etapas, são elas:

Figura 2.10 – O ciclo de vida do conteúdo digital da Digital NZ

Fonte: DigitalNZ (2006), traduzido pela autora.

1. **Selecionar**: Para conteúdo analógico ou novo, selecionar o que deve ser digitalizado e criar uma política de seleção escrita. Quando considerar a digitalização do conteúdo, é essencial avaliar se a prática é realmente necessária para aumentar o acesso, descoberta e uso do conteúdo não digital. Considerações sobre o risco para o conteúdo original, raridade,

[18]Traduzida para o português, pode ser compreendida como uma interface de programação de aplicação. Ou seja, APIs são mecanismos que permitem que dois componentes de software se comuniquem usando um conjunto de definições e protocolos.

custos de acesso, demanda e tendência de uso a longo prazo são fatores limitantes. Além disso, deve-se identificar, entender e respeitar os direitos autorais do conteúdo e incorporar essas informações aos metadados do conteúdo.

2. **Criar**: Colocar o conteúdo em uma forma utilizável. Se o conteúdo pode ser utilizado por um longo período, é importante antecipar que algumas tecnologias e padrões se tornarão obsoletos. Para isso, tecnologias de *hardware* e *software* que utilizem padrões abertos e orientações que estejam sendo mantidas e atualizadas são recomendadas.

3. **Descrever**: Descrever o conteúdo para que possa ser organizado. Para que este conteúdo seja armazenado, encontrado e usado ao longo do tempo, é essencial ter boas práticas de nomeação de arquivos e metadados associados que descrevam o que é o conteúdo, de onde veio e quem pode usá-lo.

4. **Gerir**: Gerenciar o conteúdo para mantê-lo utilizável e disponível. As boas práticas em coleta e arquivamento incluem ter uma política escrita para gerenciar o conteúdo digital (abrangendo permissão para acesso e descarte, o que pode ser coletado, padrões e formatos usados, etc.); uso de metadados administrativos; designar um responsável pela inserção de informação, atrelado a isso está ter um recurso de autenticação ou segurança que controle o acesso e as alterações; usar um sistema de classificação hierárquica e nomeação de arquivos; e identificar as etapas necessárias para fazer backup, arquivar e migrar o conteúdo para preservação.

5. **Preservar**: Gerenciar o conteúdo para mantê-lo utilizável e disponível a longo prazo. O armazenamento para preservação envolve o planejamento de cópias arquivadas do seu conteúdo para serem migradas e contingências para transferência para um novo proprietário.

6. **Descobrir**: Organizar o conteúdo para torná-lo encontrável.

7. **Usar e reutilizar**: Garantir que o conteúdo possa ser usado e reaproveitado. É incentivado o respeito às declarações de direitos e

> licenças claras que se concentrem em comportamentos permitidos, fornecendo valor que não dependa do controle de cópias, e pensar cuidadosamente sobre a colocação de conteúdo restrito online. (DIGITALNZ, 2006a)

A análise dos modelos discutidos evidencia a crescente importância e relevância das atividades de curadoria digital e de preservação e mostra que, quando devidamente planejadas e executadas em todos os estágios do gerenciamento, demonstra efetividade na preservação a longo prazo bem-sucedida dos ativos digitais. Esses elementos são cruciais para o desenvolvimento do modelo proposto neste trabalho, que aborda os desafios específicos da curadoria digital na área da saúde.

2.6 Preservação Digital

Frequentemente confundida com a curadoria digital, a preservação digital é especificamente focada na garantia de que as coleções digitais sejam acessíveis ao público no futuro. WEBB (2003) define a preservação digital com um conjunto de processos responsáveis por garantir o acesso continuado durante períodos superiores à esperança de vida do ambiente tecnológico necessário à interpretação e/ou reprodução dessa informação. Esses processos envolvem abordagens técnicas e estratégicas para definição de uma política de preservação de ativos digitais, e em qualquer organização deve abarcar os aspectos organizacionais, legais e técnicos.

A preservação digital encaixa-se como uma atividade específica, dentro do ciclo de vida da curadoria digital, que se preocupa com os meios e ações necessários para lidar com as falhas e as fragilidades das mídias, assim como com a obsolescência tecnológica (SIEBRA; SILVA, 2021). Oliver e Harvey (2016) discutem que a curadoria digital e a preservação digital têm como objeto de interesse a manutenção da informação digital ao longo do tempo. Eles destacam que, embora as definições sejam próximas, a curadoria se apresenta como evolução natural dos processos de preservação de recursos digitais.

No âmbito organizacional, é crucial definir os objetivos da instituição em relação à preservação digital, reunir uma equipe multidisciplinar responsável por essa tarefa, alocar recursos financeiros adequados e estabelecer atos administrativos que

formalizem essas ações em políticas de preservação digital. Uma política de preservação digital deve envolver todos os aspectos de um objeto digital, como criação de uma política de avaliação e seleção do material, definição de metadados, estratégias para cada classe de objeto, política de continuidade, financiamento sustentável, objetivos a nível social e organizacional, entre outros (FERREIRA, 2006).

Os aspectos legais também devem ser considerados ao definir as políticas e práticas de preservação digital. O grupo legal de uma organização deve estar amparada em leis que respaldem a instituição e garantam a propriedade intelectual ao autor do objeto digital e sua autenticidade (GRÁCIO, 2012). A Lei Geral de Proteção de Dados (LGPD), Lei nº 13.709/2018, impõe requisitos para garantir a segurança e a privacidade dos dados de saúde, considerados dados pessoais sensíveis pela lei. Isso inclui medidas para proteger contra acesso não autorizado, garantir a integridade dos dados e implementar políticas de retenção adequadas para preservar essas informações de forma segura ao longo do tempo (BRASIL, 2018a).

No mesmo contexto, o Regulamento Geral sobre a Proteção de Dados (RGPD), lei de privacidade da União Europeia (UE), estabelece que o tratamento de dados relativos à saúde constitui tratamento de categorias especiais de dados pessoais. Adicionalmente, determina que qualquer pessoa que tenha acesso a esses dados está sujeita a um dever de sigilo (UE, 2016).

A lei brasileira nº 12.527/2011 também possui uma relação indireta com as práticas relacionadas à preservação digital. Essa lei, conhecida como Lei de Acesso à Informação, estabelece o direito fundamental de acesso às informações públicas aos cidadãos e instituições. A lei exige que os órgãos públicos forneçam acesso a informações e documentos sob sua guarda. Isso implica garantir a preservação adequada dos registros digitais para atender às solicitações de acesso. Essas informações digitais precisam ainda ter garantia de integridade e autenticidade. Dessa forma, as políticas devem ser implementadas de forma a assegurar a acessibilidade e a integridade desses registros durante todo o período de retenção. (BRASIL, 2011a).

Diversos países têm desenvolvido projetos e iniciativas voltados para a preservação digital na saúde. Entre os exemplos, destacam-se o *Polo archivistico dell'Emilia-Romagna (ParER)* na Itália, responsável pela preservação permanente de registros digitais provenientes de todas as administrações públicas regionais, incluindo dados de

saúde (POLO ARCHIVISTICO DELL'EMILIA-ROMAGNA, 2014); a plataforma eBiblioCCE, criada pelo Centro Nacional de Cirurgia na Espanha; o Sistema de Preservação de Recursos Eletrônicos (SPER) dos *National Institutes of Health* nos Estados Unidos; e a metodologia de auditoria TRAC, que avalia a capacidade de sistemas de informação para manter dados digitais seguros a médio e longo prazo, também desenvolvida na Espanha. Outras inovações incluem a proposta europeia para armazenamento de dados em DNA, entre várias outras iniciativas semelhantes (MELLO; VIANNA, 2019).

No campo técnico, a preservação digital deve abordar questões como seleção e descarte de objetos digitais, adoção de modelos, padrões e iniciativas de preservação reconhecidos, a criação de metadados que permitam a identificação e recuperação eficaz dos objetos digitais, a preservação da autenticidade desses objetos, a infraestrutura tecnológica necessária, a implementação de repositórios institucionais, estratégias de preservação específicas e o suporte técnico necessário para a gestão dessas ações. Essa definição de GRÁCIO (2012) se alinha com o conceito de plano de preservação digital definido por BECKER *et al.* (2009):

> Um plano de preservação define uma série de ações de preservação a serem tomadas por uma instituição devido a um risco identificado para um determinado conjunto de objetos ou registros digitais (chamado de coleção). O Plano de Preservação leva em consideração as políticas de preservação, obrigações legais, restrições organizacionais e técnicas, requisitos do usuário e objetivos de preservação, e descreve o contexto de preservação, as estratégias de preservação avaliadas e a decisão resultante para uma estratégia, incluindo a justificativa para a decisão. Ele também especifica uma série de etapas ou ações (chamadas de plano de ação de preservação), juntamente com responsabilidades, regras e condições para execução na coleção. Desde que as ações e sua implementação, bem como o ambiente técnico, permitam isso, esse plano de ação é uma definição de fluxo de trabalho executável. (BECKER *et al.*, 2009, tradução nossa).

Em "*The Theory and Craft of Digital Preservation*", OWENS (2017) enumera ainda alguns pontos relevantes que destacam a importância de lidar com a preservação dos objetos digitais como um campo especializado e que requer uma abordagem cuidadosa e estruturada. No manuscrito, o autor descreve dezesseis axiomas que

acredita que devem servir de base para qualquer trabalho de preservação digital. OWENS (2017) enfatiza que não existe *software* que "faça" preservação digital. Um repositório para armazenar e preservar objetos digitais é o conjunto de recursos financeiros, *hardware*, tempo de equipe e implementação contínua de políticas e planejamento para garantir acesso de longo prazo. Nesse contexto, as instituições são responsáveis pela natureza contínua e colaborativa da preservação digital, pois são elas quem fornecem o ambiente, os recursos e a *expertise* necessários para proteger o patrimônio digital ao longo do tempo.

A preservação é o resultado do trabalho de pessoas e do comprometimento de recursos, não é um processo terminável, mas um processo contínuo de compreensão dos riscos e ameaças de preservação. Em muitos casos, é recomendável começar implementando ferramentas e práticas simples e então investir em processos mais complexos para continuar preservando. O equilíbrio entre esses fatores e a consideração de ameaças de preservação é a melhor forma de utilização dos recursos para proteger o conteúdo digital.

A preservação digital ainda vai além de fazer cópias de objetos ou *backups* regulares dos dados, envolve considerações mais complexas, como a sustentabilidade dos formatos de arquivo, a manutenção da integridade dos dados ao longo do tempo e a garantia de que o conteúdo permaneça acessível para as futuras gerações. É também desafiador definir o que faz parte do objeto digital e o que é exterior a ele, afinal os objetos individuais referenciam, incorporam e utilizam aspectos de outros objetos como parte de sua função diária. Em alguns casos, o conteúdo de um disco rígido pode ser gerenciado como um único item, e em outros são uma coleção de itens (OWENS, 2017).

À medida que a tecnologia muda, surgem preocupações que os métodos utilizados hoje para preservar os materiais digitais não serão suficientes ou até viáveis no futuro pois as tecnologias continuarão a progredir, tornando a geração anterior obsoleta, inevitavelmente. Decidir o que é importante sobre um objeto ou um conjunto de objetos depende em grande parte do que seu uso futuro pode ser, por isso, é fundamental, pelo menos em algum nível, pensar e estar cientes das tendências no desenvolvimento de tecnologias digitais, afinal, a obsolescência tecnológica é uma realidade incontornável. Assim, embora a preservação digital não possa eliminar os desafios da obsolescência, ela tem a capacidade de reduzir seus efeitos sobre os

objetos digitais, prevenindo, desse modo, a perda de registros de importância social. Esse fator destaca também a natureza ágil e adaptativa das políticas de preservação digital, que requer atenção constante e recursos dedicados. (OWENS, 2017)

A fim de identificar boas práticas relacionadas aos aspectos supracitados, a DPC mantém e atualiza o *Digital Preservation Policy Toolkit*[19], que também fornece informações para auxiliar a construção de uma política de preservação digital. Uma política de preservação digital expressa um conjunto de princípios que orientam uma organização na maneira como ela aborda atividades e responsabilidades de preservação. Embora uma estratégia institucional seja geralmente um plano de longo prazo para a realização de sua visão e metas, muitas organizações se referem ao seu documento de política como sua estratégia. Estratégia e política são complementares, mas geralmente tentam alcançar coisas diferentes.

Uma política de preservação digital deve se relacionar com os objetivos organizacionais e outras políticas; estabelecer e ser facilmente compreendida por seu público-alvo; deve identificar papéis, responsabilidades e escopo; e deve ser revisada regularmente em conformidade com outra documentação organizacional (DIGITAL PRESERVATION COALITION, 2023). A política de preservação estabelece os padrões e procedimentos específicos para alcançar esses objetivos. Entre os padrões mais populares no campo da preservação está o OAIS, um modelo conceitual que visa identificar os componentes funcionais que deverão fazer parte de um sistema de informação dedicado à preservação digital, bem como as suas interfaces internas e externas e os objetos de informação trocados no seu interior (FERREIRA; SARAIVA; RODRIGUES, 2012).

---

[19] O toolkit pode ser encontrado em: https://www.dpconline.org/digipres/implement-digipres/policy-toolkit

2.6.1 O modelo de referência Open Archival Information System (OAIS)

Como um esforço da *Consultative Committee for Space Data Systems* (CCSDS) para desenvolver padrões formais para o armazenamento de longo prazo de dados digitais gerados a partir de missões espaciais, um modelo de referência para um "sistema de informação arquivística aberto" foi desenvolvido. Em 2003, o modelo foi aprovado como uma *International Organization for Standardization* (ISO), e revisado pela norma ISO 14721:2012[20]. O modelo de referência OAIS foi desenvolvido por meio de um processo aberto e iterativo de elaboração, análise e revisão; *feedback* da comunidade foi fornecido por meio de discussões em *workshops* presenciais e respostas por escrito a solicitações formais de comentários. (LAVOIE, 2014)

O conceito principal do modelo de referência é o de um *open archival information system* (OAIS). O termo *open* (aberto) não diz respeito ao nível de acessibilidade associado a um arquivo, mas refere-se ao fato de que o modelo de referência foi desenvolvido e lançado em fóruns públicos abertos, nos quais qualquer parte interessada foi incentivada a participar. Um *archival information system* (sistema de informação arquivística) é "uma organização, que pode ser parte de uma parte de uma organização maior, de pessoas e sistemas que aceitaram a responsabilidade de preservar informações e disponibilizá-las para um público-alvo específico".

A primeira responsabilidade de um OAIS é estabelecer critérios de seleção explícitos para determinar quais materiais são apropriados para inclusão no depósito de arquivos. A segunda responsabilidade enfatiza a necessidade de um OAIS obter direitos de propriedade intelectual suficientes, em conjunto com a custódia dos itens, para autorizar os procedimentos necessários e atender os objetivos de preservação. Também é necessário determinar a comunidade-alvo primária, pois muitas vezes para que as informações preservadas sejam compreendidas, o contexto em que estão inseridas também precisa ser entendido.

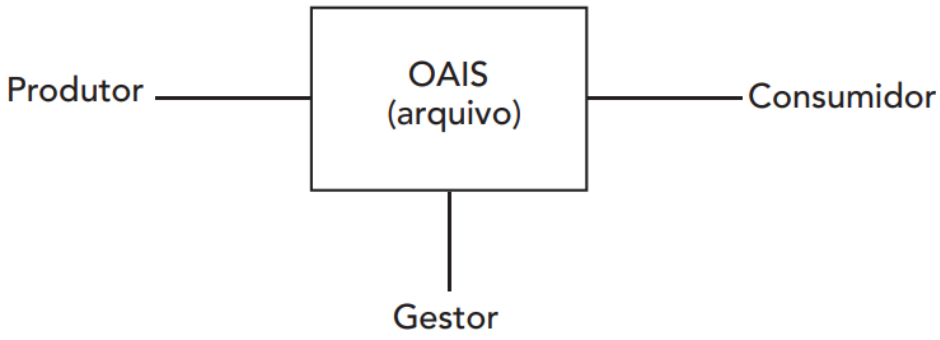
---

[20] ISO 14721:2012: https://www.iso.org/standard/57284.html

Diante disso, o OAIS deve não apenas preservar informações, mas também uma quantidade suficiente de seu contexto associado para garantir que as informações sejam compreensíveis e utilizáveis pelas gerações futuras. Um OAIS também deve estabelecer e documentar políticas e procedimentos claros para realizar a preservação das informações; e, finalmente deve ter o dever de disponibilizar o conteúdo de seu acervo arquivístico à comunidade de usuários a que se destina, por meio da implementação de mecanismos e serviços de acesso que atendam, na medida do possível, às necessidades e exigências dos usuários. (LAVOIE, 2014)

O ambiente de um OAIS é composto por quatro componentes distintos, três dos quais são explicitamente externos ao OAIS: Gestor (*Management*), Produtor (*Producer*) e Consumidor (*Consumer*) (Figura 2.11). O Gestor é responsável por atividades que incluem o planejamento estratégico, definição do escopo do OAIS, e a negociação da garantia de preservação associada aos itens confiáveis do sistema. Os Produtores são os indivíduos, organizações ou sistemas que fornecem as informações a serem preservadas, assim como o conteúdo e metadados associados. Os Consumidores são os indivíduos, organizações ou sistemas que interagem e consomem as informações designadas para preservação no OAIS (GRÁCIO, 2012).

Figura 2.11 – O ambiente OAIS



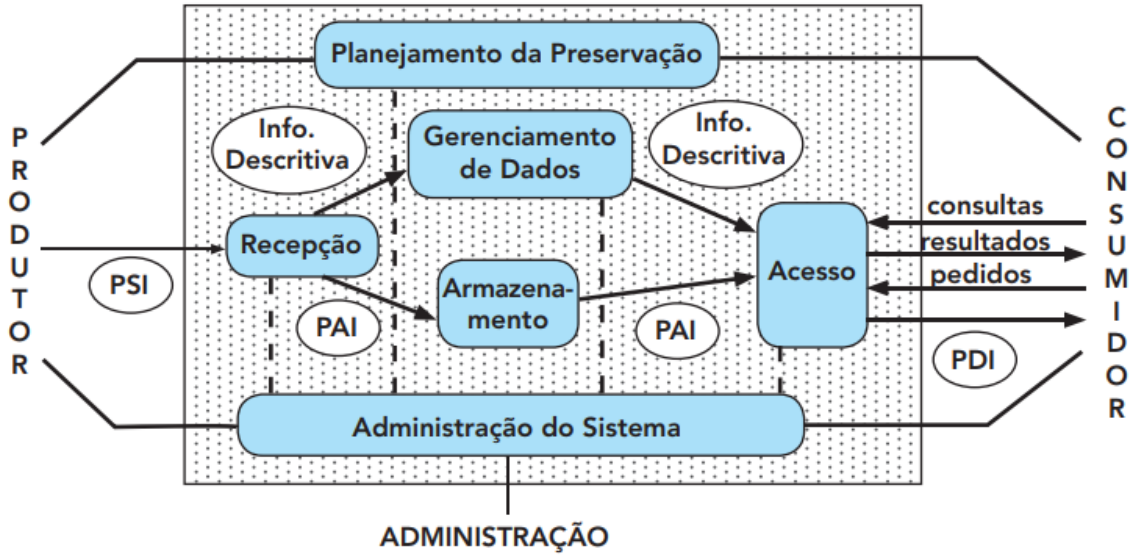
Fonte: GRÁCIO (2012)

O CCSDS (2012) destaca que, embora não esteja explicitamente representado na Figura, deve-se observar que o ambiente externo de um OAIS também pode incluir interação com outros OAIS. Os componentes do ambiente representam papeis funcionais em vez de papeis organizacionais, o ponto-chave é a separação lógica dos papeis de tomada de decisão e das partes interessadas ligadas às atividades de preservação digital. O modelo de referência identifica e descreve os mecanismos

principais para que um OAIS consiga preservar informações a longo prazo e disponibilizá-las para os potenciais Consumidores.

Esses mecanismos são resumidos pelo modelo funcional do OAIS (Figura 2.12): uma coleção de seis entidades funcionais (destacadas em azul) que, juntas, cumprem o papel de preservar e fornecer acesso às informações de um OAIS. Essas entidades podem ser implementadas e configuradas de qualquer maneira apropriada às circunstâncias e tecnologia específicas. (CCSDS, 2012)

Figura 2.12 – Modelo de referência OAIS



Fonte: Grácio (2012)

As funções da entidade funcional de **Recepção** (*Ingest*) incluem receber Pacote de Informação de Submissão (PSI)[21], realizar garantia de qualidade, gerar Pacote de Informação de Arquivo (PIA)[22] que esteja em conformidade com os padrões de

---

[21] *Submission Information Packages* (SIP), em inglês. É o pacote enviado a um OAIS por um Produtor, geralmente possui o conteúdo da informação (*content information*) (objeto físico ou digital e descrição que possibilita a representação da informação) e alguns Pacote de Informação de Disseminação (PID) (informações necessárias do conteúdo da informação, que incluem sua origem, histórico, relacionamento com outras informações, identificadores, termos de acesso etc.).

[22] *Archival Information Package* (AIP), em inglês. Consiste em um conjunto completo de PDI para o conteúdo da informação associado. Dentro do OAIS, um ou mais PSIs são transformados em um ou mais PAIs.

formatação e documentação de dados do Armazenamento, extrair Informações Descritivas dos PIAs para inclusão no banco de dados e coordenar atualizações para Armazenamento Arquivístico e Gerenciamento de Dados. O **Armazenamento** (*Archival Storage*) fornece os serviços e funções para o armazenamento, manutenção e recuperação de PAIs recebidas da Recepção. As funções da entidade englobam a adição de PAIs ao armazenamento permanente, gerenciar a hierarquia de armazenamento, atualizar a mídia de armazenamento, realizar procedimentos de verificação de erros e de recuperação de desastres, e fornecer PAIs para a entidade funcional Acesso.

O **Gerenciamento de Dados** (*Data Management*) é responsável por manter bancos de dados de metadados descritivos e administrativos. Isso inclui identificar e descrever as informações arquivadas e apoiar operações internas, como consultas e geração de relatórios, executar atualizações na base de dados. A entidade funcional **Planejamento da Preservação** (*Preservation Planning*) é responsável por elaborar a estratégia de preservação e recomendar revisões conforme as condições do ambiente OAIS evoluem. Essa função monitora o ambiente externo e identifica mudanças e riscos que possam afetar a capacidade do OAIS de preservar e garantir o acesso às informações, e, com base nisso, desenvolve recomendações para atualizar as políticas e procedimentos.

O **Acesso** (*Access*), como o próprio nome sugere, gerencia os processos e serviços pelos quais os Consumidores localizam, solicitam e recebem itens armazenados no OAIS. A entidade processa as consultas sobre os itens armazenados e coordena a recuperação e entrega de conteúdo solicitado. O Acesso também implementa mecanismos de segurança e controle de acesso associados ao conteúdo arquivado. o acesso aos objetos digitais é feito através do PID, que se diferencia do PIA, apesar de ambos terem origem no mesmo objeto digital preservado. Essa diferença na nomenclatura sublinha que a informação entregue ao *Consumidor* pode ser adaptada em termos de forma ou conteúdo em relação àquela que é mantida no arquivo (LAVOIE, 2014).

Como destacado por Lira e Siebra (2022), as diferenças entre o PID e o PIA podem envolver o formato do conteúdo, a quantidade de dados e os metadados fornecidos ao Consumidor. Por exemplo, uma imagem armazenada em formato TIFF para preservação pode ser convertida para JPEG para facilitar a disseminação, garantindo

um acesso mais eficiente e compatível com os dispositivos ou sistemas utilizados pelo *Consumidor*. Além disso, o PID pode reunir informações de um ou mais PIAs, ou até mesmo conter apenas uma parte de um PIA, adaptando a quantidade de conteúdo entregue conforme as necessidades do usuário.

As autoras complementam que, de maneira similar, o PID fornece uma quantidade reduzida de metadados ao *Consumidor*, em comparação ao conjunto completo presente no PIA. Enquanto o PIA inclui metadados de preservação, técnicos e administrativos, essenciais para a integridade a longo prazo do objeto digital, o PID se limita aos metadados relevantes para o usuário, excluindo informações que não têm utilidade prática para ele. Assim, o PID proporciona uma experiência mais simplificada ao usuário, equilibrando praticidade de acesso com a preservação dos dados, e é fundamental para disponibilizar os dados de forma adequada aos objetivos de disseminação.

A entidade funcional **Administração do Sistema** (*Administration*) é responsável por gerenciar as operações diárias do OAIS, coordenar as atividades das outras cinco entidades funcionais e interagir com Produtores, Consumidores e Gerenciamento (LAVOIE, 2014). O OAIS é um modelo conceitual que não especifica tecnologias ou arquiteturas de sistema, mas várias iniciativas têm usado seus conceitos como base para sistemas de arquivamento funcionais.

2.6.2 O modelo de preservação Hipátia

O modelo Hipátia, proposto pelo IBICT, propõe uma abordagem para a preservação digital arquivística que integra ambientes de gestão, preservação e acesso, assegurando a autenticidade dos documentos ao longo do tempo. A estrutura do modelo fundamenta-se em normativas tanto nacionais quanto internacionais, focadas nas melhores práticas de preservação digital e que abordam também aspectos de gestão e acesso (IBICT, [s.d.]).

A aplicação do modelo Hipátia é organizada em cinco etapas principais: preparação arquivística, preparação computacional, extração de objetos digitais, preservação e disseminação. Essas fases podem ocorrer de forma simultânea, como nas etapas de preparação arquivística e computacional, ou de maneira sequencial, como nas etapas

de extração, preservação e disseminação de objetos digitais. O modelo, visto na Figura 2.13, é geralmente representado em formato sequencial para facilitar a compreensão das interdependências e do fluxo de trabalho entre as etapas (BRAGA; ARELLANO, 2022).

Figura 2.13 – Etapas do Modelo Hipátia

Fonte: (BRAGA; ARELLANO, 2022)

Como descrito por BRAGA; ARELLANO (2022), a primeira fase, a *Preparação arquivística*, define quais objetos digitais serão preservados, identificando dados, metadados, e a estrutura de cada objeto digital. Essa etapa cria as diretrizes essenciais para todas as fases seguintes e envolve a avaliação documental e análise de temporalidade dos dados. Na *Preparação computacional*, o ambiente de tecnologia necessário é configurado para implementar o modelo, incluindo a arquitetura de rede, permissões de acesso e a instalação dos sistemas necessários. Essa etapa depende da colaboração entre profissionais de arquivologia e tecnologia.

Na *Extração de objetos digitais*, o sistema produtor é conectado ao ambiente de preservação, usando APIs, bancos de dados, ou arquivos do sistema operacional. Os dados e metadados são embalados em pacotes, que garantem segurança e fácil acesso hierárquico. A extração pode ocorrer de forma automatizada por intervalos de tempo, sob demanda, ou por eventos. Na fase de *Preservação*: Os objetos digitais são transferidos para o repositório (Archivematica – tratado no tópico 2.6.5 Sistemas de Preservação) que segue o modelo OAIS. Pacotes de informações (PSI, PIA e PID) são gerados para garantir a preservação a longo prazo. Essa etapa organiza os objetos para futuras consultas e garante a integridade e autenticidade dos dados.

Por fim, na *disseminação*, os objetos preservados são disponibilizados para o público conforme políticas institucionais e regulatórias, incluindo a LGPD. O sistema AtoM (tratado no tópico 2.6.5 Sistemas de Preservação) é utilizado para acesso público. A implementação desse modelo depende de soluções tecnológicas integradas, três principais ferramentas foram desenvolvidas com essa finalidade: BarraPres, ValidaPres e MetaPres.

O BarraPres é um barramento que atua como uma interface entre os sistemas gestores de conteúdo e o sistema de preservação, padronizando e organizando os pacotes de dados para preservação. Desenvolvido pelo IBICT como software livre, o BarraPres integra o modelo OAIS, convertendo os conteúdos recebidos para o formato de pacotes BagIt[23], que são utilizados pelo Archivematica. Ele cria PSI que, junto com metadados específicos de preservação, são enviados ao repositório de preservação, garantindo que os documentos digitais e seus metadados sejam armazenados de maneira segura e organizada (IBICT, [s.d.]).

Conforme descrição no *website* do IBICT, a plataforma ValidaPres é responsável por conceder e validar a autenticidade dos documentos digitalizados, permitindo que os documentos recebam elementos descritivos que garantem a credibilidade e autenticidade a longo prazo e permite sua homologação por meio de assinaturas digitais simples ou qualificadas., auxiliando na governança e no gerenciamento de projetos de digitalização. Além disso, ele padroniza metadados e certifica a qualidade da imagem, selecionando formatos e compressões que atendam à legislação. O MetaPres é uma solução de gerenciamento de metadados que utiliza algoritmos em *Python* para operar com repositórios arquivísticos digitais que seguem o modelo Hipátia. Ele suporta o uso do padrão ISAD(G) para metadados descritivos. Existe expectativa para que, no futuro, o MetaPres seja expandido para disseminar PIDs em repositórios do modelo Hipátia com outros *softwares* de acesso e disseminação.

Paralelamente a essas soluções, os Níveis de Preservação Digital, desenvolvidos pela National Digital Stewardship Alliance (NDSA)[24], representam outra contribuição significativa para a preservação digital. A NDSA foi criada pela Biblioteca do Congresso dos EUA em 2010, e seus níveis consistem em um conjunto de diretrizes que auxiliam organizações a avaliar e aprimorar suas práticas de preservação digital, estruturados em quatro estágios progressivos de maturidade (BUARQUE; MACHADO; PONTES, 2020). Além dos próprios níveis, a NDSA desenvolveu uma

---

[23] O BagIt é um formato de empacotamento que usa uma estrutura de pastas para organizar e transferir arquivos digitais com segurança, incluindo seus metadados. Foi criado pela Biblioteca do Congresso dos EUA.

[24] Disponível em: https://ndsa.org/

ferramenta adicional conhecida como matriz dos Níveis de Preservação Digital. Essa matriz permite uma visualização simplificada de quais critérios estão associados a cada nível de preservação. Na próxima seção, exploraremos como essas ferramentas se relacionam e sua relevância para o contexto da preservação digital.

2.6.3 Níveis de Preservação Digital da National Digital Stewardship Alliance (NDSA)

Os Níveis da Preservação Digital são um conjunto de diretrizes e práticas para ajudar os profissionais de preservação digital a criarem ou avaliarem seu programa de preservação digital. Originalmente criada em 2013, a segunda versão foi lançada em 2019 com um conjunto de documentos e recursos adicionais (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022). A Matriz dos Níveis de Preservação Digital está representada na Tabela 2.3.

Tabela 2.3 – Matriz dos Níveis de Preservação Digital

Área Funcional	**Nível**			
	Nível 1 (Conheça o seu conteúdo)	**Nível 2 (Proteja o seu conteúdo)**	**Nível 3 (Supervisione o seu conteúdo)**	**Nível 4 (Mantenha seu conteúdo)**
Armazenamento	Ter duas cópias completas em locais separados	Ter três cópias completas, com pelo menos uma cópia em um local geograficamente separado	Ter pelo menos uma cópia em um local geográfico com uma ameaça de desastre diferente das outras cópias	Ter pelo menos três cópias em locais geográficos com uma ameaça de desastre diferente das outras cópias
	Documentar todas as	Documentar o armazenament	Ter pelo menos uma	Maximizar a diversificação

	mídias de armazenamento onde o conteúdo está armazenado	o e a mídia de armazenamento indicando os recursos e as dependências necessários para o seu funcionamento	cópia em um tipo diferente de mídia de armazenamento	das mídias de armazenamento para evitar pontos únicos de falha
	Armazenar o conteúdo em uma mídia de armazenamento estável		Monitorar a obsolescência do armazenamento e da mídia	Ter um plano e executar ações para lidar com a obsolescência de hardware, software e mídia de armazenamento
Integridade	Verificar as informações de integridade se elas tiverem sido fornecidas com o conteúdo	Verificar as informações de integridade ao mover ou copiar conteúdo	Verificar as informações de integridade do conteúdo em intervalos fixos	Verificar informações de integridade em resposta a eventos ou atividades específicas
	Gerar informações de integridade se não forem fornecidas com o conteúdo	Utilizar bloqueadores de escrita ao trabalhar com a mídia original	Documentar os processos e os resultados da verificação das informações de integridade	Substituir ou reparar conteúdo corrompido, conforme necessário
	Verificar vírus	Fazer backup	Realizar	

	em todo o conteúdo; isolar o conteúdo para quarentena, conforme necessário	das informações de integridade e armazenar a cópia em um local separado do conteúdo	auditoria das informações de integridade sob demanda	
Controle	Determinar os agentes humanos e de software que devem ser autorizados a ler, gravar, mover e excluir conteúdo	Documentar e aplicar as informações sobre os agentes humanos e de software autorizados a ler, gravar, mover e excluir conteúdo	Manter registros e identificar os agentes humanos e de software que realizaram ações no conteúdo	Realizar revisão periódica de ações/registros de acesso
Metadados	Criar inventário de conteúdo, documentando também os locais de armazenamento atuais	Armazenar metadados suficientes para saber o que é o conteúdo (isso pode incluir alguma combinação de metadados administrativos, técnicos, descritivos, de preservação e estruturais)	Determinar quais padrões de metadados devem ser aplicados	Registrar ações de preservação associadas ao conteúdo e quando essas ações ocorrem
	Fazer backup do inventário e armazenar pelo menos uma cópia separadamente do conteúdo		Encontrar e preencher lacunas nos metadados para atender a esses padrões	Implementar os padrões de metadados escolhidos

Conteúdo	Documentar os formatos de arquivos e outras características essenciais de conteúdo, incluindo como e quando foram identificados	Verificar formatos de arquivos e outras características essenciais do conteúdo Estabelecer relações com criadores de conteúdo para incentivar escolhas sustentáveis de arquivos	Monitorar a obsolescência e as mudanças nas tecnologias das quais o conteúdo depende	Realizar migrações, normalizações, emulação e atividades semelhantes que garantam que o conteúdo possa ser acessado

Fonte: Levels of Preservation Revisions Working Group (2023), traduzido pela autora

A matriz apresentada organiza práticas recomendadas em quatro níveis de maturidade através de cinco áreas funcionais: Armazenamento, Integridade, Controle, Metadados e Conteúdo. No Nível 1, a ênfase está em estabelecer uma base para a preservação digital, com ações como manter duas cópias completas do conteúdo em locais separados, documentar os meios de armazenamento e garantir que o conteúdo esteja em armazenamento estável. Este nível também envolve a verificação de integridade se esta for fornecida, gerar informações de integridade se não estiverem disponíveis, e realizar verificações de vírus.

Já no Nível 2, o foco é em fortalecer a proteção do conteúdo, incluindo a manutenção de três cópias completas com pelo menos uma em uma localização geográfica diferente, o uso de bloqueadores de gravação ao trabalhar com o conteúdo original, e o backup das informações de integridade em locais separados. Também abrange a documentação de agentes autorizados a acessar e modificar o conteúdo, e o armazenamento seguro de inventários de conteúdo.

Com o objetivo de garantir a continuidade e a integridade do conteúdo ao longo do tempo, o nível 3 inclui manter pelo menos uma cópia em um local com uma ameaça

de desastre diferente, rastrear a obsolescência dos meios de armazenamento e realizar verificações de integridade em intervalos fixos. A documentação dos processos de verificação e a manutenção de *logs* de ações e acessos também são cruciais neste nível.

O último nível, o Nível 4, a ênfase está na sustentabilidade a longo prazo do conteúdo. Este nível requer a maximização da diversidade de armazenamento para evitar pontos únicos de falha, a implementação de ações para resolver a obsolescência de *hardware*, *software* e mídia, e a substituição ou reparo de conteúdo corrompido conforme necessário. Além disso, envolve a revisão periódica dos *logs* de ações e acessos, a aplicação de normas de metadados, e a realização de migrações, normalizações e emulações para garantir o acesso contínuo ao conteúdo.

Embora útil para avaliar o estágio de preservação digital em uma instituição e identificar lacunas, o modelo enfatiza estratégias específicas de preservação e incentiva sua aplicação, que servem mais como um guia prático para a implementação dessas atividades.

2.6.4 Estratégias de Preservação

A obsolescência tecnológica tem sido considerada um desafio significante para a preservação digital a longo prazo. De forma simples, a obsolescência é o processo de tornar-se obsoleto ou sem utilidade (DPC, 2023). A DPC também aponta que, ao falarmos sobre obsolescência tecnológica, podemos considerar a "obsolescência institucional", ou seja, quando a tecnologia em questão não está mais em uso ou facilmente acessível por uma determinada instituição. Isso torna-se um problema quando compromete o significado do conteúdo ou sua interpretação por um usuário.

Um dos principais objetivos das estratégias de preservação digital é manter a confiabilidade e a autenticidade do material preservado, apesar dessas mudanças geracionais na tecnologia da computação. A autenticidade dos objetos digitais é ameaçada sempre que eles são transferidos ao longo do tempo ou entre diferentes sistemas. Para garantir que a autenticidade seja mantida, é essencial definir claramente a identidade dos materiais e proteger sua integridade ao longo do tempo. A confiabilidade, por sua vez, refere-se à credibilidade do conteúdo digital, ou seja, à

sua capacidade de ser considerado uma representação fiel e precisa de um fato ou informação (INTERPARES PROJECT, 2011).

Para garantir esses critérios, os Níveis da Preservação Digital por exemplo, estabelecem práticas para garantir a preservação do conteúdo digital a longo prazo, como ilustrado na Tabela 2.3. Além dessas estratégias, é possível explorar outras técnicas e tecnologias que podem ser adotadas e combinadas entre si para preservar objetos digitais. As estratégias operacionais de preservação digital dizem respeito ao que de fato pode ser feito, as técnicas e tecnologias utilizadas para preservar objetos digitais (MOREIRA, 2017). Entre essas estratégias estão: Conservação/Preservação da tecnologia, Migração, Emulação, Encapsulamento e Refrescamento.

REFRESCAMENTO

De acordo com Ferreira (2006), o refrescamento de suporte consiste na transferência de informações de uma mídia de armazenamento para outra mais atual, antes que a mídia original se deteriore ou se torne obsoleta. O autor também ressalta que o refrescamento, por si só, não constitui uma estratégia de preservação, mas é um pré-requisito essencial para o sucesso de qualquer estratégia de preservação de longo prazo. Assim, o refrescamento periódico das mídias, juntamente com a verificação frequente de sua integridade, são atividades que devem fazer parte desses processos

PRESERVAÇÃO DA TECNOLOGIA

A preservação da tecnologia é uma estratégia que propõe a conservação do contexto tecnológico utilizado originalmente na criação dos objetos digitais que se pretende preservar (FERREIRA, 2006). Mas, para tanto, ainda de acordo com FERREIRA (2011), manter o acesso a essas tecnologias "implica a criação de verdadeiros museus de *software* e *hardware* obsoletos". Portanto, a longo prazo, essa estratégia exige investimentos significativos em espaço físico, manutenção técnica e recursos financeiros, o que a torna inviável. Além disso, o acesso à determinado conteúdo fica dependente e restrito apenas ao local físico onde estão os *hardwares* preservados.

No que diz respeito à conservação de *software*, os recursos necessários incluem profissionais especialistas e ferramentas tecnológicas específicas. Existe ainda uma incerteza legal nessa prática. As estruturas legais criadas para regulamentar o *software* no mercado, associados a licenças e outras restrições de propriedade intelectual. muitas vezes não acompanham as necessidades de preservação, o que

pode dificultar ou até mesmo impedir ações de conservação em certos contextos (ARL; CMSI; PIJIP, 2018). A conservação do *software* trata-se de uma estratégia recente, e é a combinação de duas estratégias: o Encapsulamento e a Emulação (BAGGIO; FLORES, 2015), tratadas a seguir.

ENCAPSULAMENTO E EMULAÇÃO

Baggio e Flores (2015) definem que o encapsulamento é uma estratégia de preservação que consiste em preservar todos os detalhes de como interpretar o objeto digital. Preserva-se juntamente com o objeto digital, toda a informação (descrição formal e detalhada do ambiente de *software* e *hardware* requerido para seu funcionamento) necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores e ou emuladores. A emulação precisa do desenvolvimento de técnicas de encapsulamento de documentos, seus metadados, *software* e especificações de emulador de forma a assegurar sua coesão e prevenir sua corrupção. (ARELLANO, 2008)

O objetivo da emulação é preservar a aparência e a funcionalidade do objeto digital. Essa estratégia envolve o uso de um emulador, que é um software capaz de reproduzir o comportamento de uma plataforma de *hardware* e/ou *software* em outra plataforma que, inicialmente, seria incompatível. No entanto, como FERREIRA (2006) pontua, o próprio emulador poderá sofrer de obsolescência, havendo então a necessidade de convertê-lo para uma nova plataforma ou desenvolver um novo emulador capaz de emular o primeiro.

Para ARELLANO (2008), a emulação deve ser usada nos casos em que os recursos digitais não podem ser migrados dada a sua complexidade, nem convertidos para formatos de *software* independentes. A implementação dessa estratégia é complexa e difícil, devido principalmente, ao tipo de necessidades a serem preenchidas.

MIGRAÇÃO

Para Martin e Coleman (2002), no ambiente tecnológico, todos os dados digitais devem ser migrados a cada ano para que possam sobreviver. A migração está centrada na preservação do conteúdo intelectual do objeto digital e consiste na transferência de materiais digitais de uma plataforma computacional, *hardware* e *software*, em vias de descontinuidade, para outra mais moderna (FERREIRA, 2006).

No entanto, nos processos de migração, há uma probabilidade significativa de que algumas das propriedades que compõem os objetos digitais não sejam transferidas corretamente para o formato de destino escolhido. Por isso, ARELLANO (2008) acrescenta que os metadados têm um papel importante em qualquer estratégia de migração bem-sucedida. Esse tipo de estratégia dependerá dos metadados criados para registrar a história da migração de um objeto digital. Também existe a necessidade de informação do contexto para ser registrada (e preservada) para que, dessa maneira, futuros usuários possam entender o ambiente tecnológico no qual um objeto digital foi criado.

Ainda é importante destacar que o formato de destino ainda se encontra sob constante ameaça de se tornar obsoleto e, portanto, é inevitável que uma nova migração tenha que ser administrada.

2.6.5 Sistemas de Preservação

Para auxiliar na execução dessas estratégias, sistemas para preservação digital e o gerenciamento de arquivos têm um papel importante de suporte. Entre eles está o Archivematica, um sistema de preservação digital gratuito e de código aberto, projetado para garantir o acesso de longo prazo a coleções de objetos digitais com base em padrões. Esse sistema é integrado ao AtoM (Access to Memory), ambos desenvolvidos pela *Artefactual Systems*, e oferece um conjunto completo de ferramentas que permitem aos usuários processarem objetos digitais desde a ingestão até o armazenamento e o acesso, em conformidade com o modelo de referência OAIS (ARTEFACTUAL SYSTEMS INC, 2024).

No caso do AtoM, o seu objetivo principal é permitir que instituições como arquivos, bibliotecas e museus cataloguem, descrevam e organizem suas coleções digitais de forma padronizada, utilizando normas arquivísticas internacionais. Ele também permite que esses registros sejam acessíveis ao público pela internet, sendo útil para a apresentação dos arquivos em uma plataforma *web*. Por outro lado, o Archivematica é uma ferramenta focada na preservação digital. Seu objetivo é oferecer aos profissionais com recursos técnicos e financeiros limitados as ferramentas, metodologias e confiança necessárias para iniciar a preservação de

informações digitais de imediato. Para isso, o projeto organiza o processo em passos concretos e específicos que devem ser executados para estar em conformidade com o modelo OAIS, desde a Ingestão até o Acesso. (ARTEFACTUAL SYSTEMS INC, 2024).

Em 2021, pretendendo ampliar o uso dessas ferramentas no Brasil, o Arquivo Nacional produziu tutoriais sobre os *softwares*, com um passo a passo sobre a utilização das ferramentas complementares de preservação e acesso à informação em ambiente digital (ARQUIVO NACIONAL, 2021). Tanto o AtoM quanto o Archivematica são desenvolvidos com ferramentas de código aberto, com documentação aberta e fóruns públicos para usuários trocarem informações e receberem suporte. Além disso, o código-fonte também está disponível gratuitamente, o que permite que desenvolvedores e usuários adaptem os *softwares* conforme as necessidades institucionais.

Outros sistemas de preservação digital se descrevem como baseados no modelo de referência do OAIS. O *The Lots of Copies Keeps Stuff Safe* (LOCKSS), por exemplo, que fornece tecnologia e soluções de preservação digital e é um componente fundamental do portfólio de bibliotecas digitais da Universidade de Stanford, produziu uma declaração formal de conformidade com a ISO 14721:2003[25] (LAVOIE, 2014). O programa foi adotado pelo Brasil desde 2002, no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e representa uma contribuição significativa para a informação científica no Brasil, que, por conseguinte, preserva também o conteúdo de publicações internacionais de grandes instituições participantes da Iniciativa LOCKSS (IBICT, 2022).

O LOCKSS é um *software* de código aberto que opera em um sistema *peer-to-peer*[26] para preservar documentos eletrônicos. Instituições participantes armazenam cópias

---

[25]Acesso em: <https://assets.lockss.org/documentation/Formal_statement_Of_Conformance_to_ISO_14721-2001.pdf>

[26] Um sistema "*peer-to-peer*" (P2P) é uma rede descentralizada onde os computadores (chamados "*peers*") compartilham recursos e se comunicam diretamente entre si, sem a necessidade de um servidor central.

autorizadas de conteúdo *online* em servidores locais, chamados "caixas LOCKSS", que reproduzem fielmente os materiais originais. O *software* coleta conteúdo usando um rastreador de Internet, verifica periodicamente a integridade das informações, e corrige quaisquer inconsistências. Ele também oferece acesso ao conteúdo preservado para usuários autorizados, através de servidores *web* e outros padrões de acesso. A interface administrativa permite a seleção e monitoramento de novos conteúdos, garantindo a preservação e a migração de formatos conforme necessário. As revistas podem autorizar o arquivamento digital configurando um "manifesto" que permite ao LOCKSS coletar e preservar seu conteúdo (MÁRDERO ARELLANO, 2014; LOCKSS, 1999).

Outro *software* de preservação, utilizado no projeto Repositório Clínico Digital (RCD) do Centro Hospitalar Universitário de São João (CHUSJ) em Portugal, é o RODA. O objetivo do RCD foi digitalizar e centralizar uma grande parte do arquivo físico do hospital em um repositório digital unificado, permitindo que os profissionais de saúde tenham acesso rápido e estruturado aos registros clínicos essenciais para suas atividades diárias. Esse projeto também envolveu a criação de um fluxo de trabalho para digitalizar e arquivar os documentos recentes, alimentando automaticamente o repositório clínico digital (FERNANDES, 2021).

Para a implementação do RCD, o RODA foi escolhido por ser uma solução de repositório digital que permite armazenar, preservar e dar acesso aos processos clínicos de milhares de pacientes e operacionais de saúde. A versão *Community* do *software* é gratuito, de código aberto, e compatível com padrões como o OAIS, Dublin Core, Encoded Archival Description (EAD)[27] e PREMIS. Além de validar pacotes de informações, o RODA realiza verificações de vírus, identifica e categoriza formatos de arquivos, extrai metadados técnicos e converte documentos para formatos mais adequados à preservação de longo prazo (KEEP SOLUTIONS, 2024). No caso do CHUSJ, a KEEP SOLUTIONS, empresa responsável pelo desenvolvimento do RODA, foi contratada para gerenciar o sistema de preservação

---

[27] O EAD é um padrão de codificação utilizado para descrever e fornecer acesso a coleções de arquivos e documentos arquivísticos utilizando a linguagem XML mantida pela Biblioteca do Congresso e pela *Society of American Archivists*. Disponível em: <https://www.loc.gov/ead/>

digital, oferecendo versões empresariais pagas e suporte personalizado (FERNANDES, 2021).

2.7 Preservação Digital no Brasil

No Brasil, o IBICT tem promovido ações e iniciativas de curadoria e preservação digital, com políticas, programas e tecnologias para garantir a longevidade e acessibilidade das informações digitais. Desde 2002, a instituição brasileira é ativa na disseminação de conhecimento e capacitação em preservação digital, oferecendo cursos e desenvolvendo políticas e estratégias de preservação. Publicações científicas e a revista "Ciência da Informação" – vinculada ao IBICT e que conta com mais de 25 publicações nas áreas – são importantes fontes de disseminação sobre o tema.

Além do Modelo Hipátia, outra principal contribuição é o projeto da Rede Brasileira de Preservação Brasileira – Cariniana, que, em parceria com instituições de ensino e pesquisa e com a colaboração de especialistas brasileiros, propõe a construção e customização de uma rede nacional de serviços de preservação digital. A Rede Cariniana tem como objetivo salvaguardar os registros da ciência, tecnologia e do patrimônio cultural do Brasil. Ela oferece uma série de alternativas para que as instituições brasileiras possam colecionar, armazenar e promover o acesso ao conteúdo selecionado através de cópias autorizadas. A rede oferece ainda pacotes de *softwares*, aplicações e ambientes multimídia para a implementação e desenvolvimento de documentos digitais preserváveis (MÁRDERO ARELLANO, 2014).

Organizado pela Rede Cariniana, o Seminário Internacional de Preservação Digital (SINPRED) também entra como um evento de grande importância para a preservação digital no Brasil e na América Latina. O SINPRED apresenta discussões sobre metodologias aplicáveis de preservação digital com o objetivo de promover o intercâmbio de informações e experiências, de modo a fortalecer a preservação digital em todas as regiões do país, difundindo o uso de soluções e tecnologias, bem como a discussão de novos serviços que possam ser oferecidos. O evento é bianual e é uma fonte importante de informações e *insights* que podem ajudar a identificar tendências e desafios nas áreas de curadoria e preservação digital no Brasil.

No Brasil, além do IBICT, o Conselho Nacional de Arquivos (CONARQ) apresenta preocupações no que concerne a preservação digital no país. O órgão é responsável pela edição de decretos regulamentadores da Lei n. 8.159[28] (BRASIL, 1991), e de resoluções que tratam de temas diversos relativos à gestão de documentos convencionais e digitais. O CONARQ publicou em 2004 a Carta para a Preservação do Patrimônio Arquivístico Digital, cujo lema é "preservar para garantir o acesso" e foi inspirada nas recomendações da United Nations Educational, Scientific and Cultural Organization (UNESCO) no documento *Guidelines for the Preservation of Digital Heritage*[29] (2003).

O documento em questão estabelece diretrizes e princípios para a preservação do patrimônio arquivístico digital no Brasil, e destaca a importância da preservação de documentos digitais, reconhecendo que o uso crescente de tecnologias digitais na administração pública e em outras áreas resulta na criação de um grande volume de registros eletrônicos que precisam ser adequadamente preservados para garantir sua autenticidade, integridade e acessibilidade ao longo do tempo. (CONARQ, 2004)

O projeto *International Research on Permanent Authentic Records in Electronic Systems* (InterPARES)[30], sediado na Universidade de Colúmbia Britânica, no Canadá, também contribuiu significativamente para o desenvolvimento de conhecimentos teóricos e práticos no Brasil relacionados à preservação de documentos arquivísticos digitais autênticos e à mitigação de riscos associados à perda de acesso a esses registros. Ao longo de suas várias fases, o projeto identificou requisitos conceituais, desenvolveu modelos conceituais, diretrizes e estratégias para preservação, e aplicou esse conhecimento em estudos de casos em parceria com instituições de todo o mundo. No contexto brasileiro, a participação do "TEAM Brasil" no projeto, coordenado pelo Arquivo Nacional e com o envolvimento de diversas instituições, trouxe contribuições valiosas para a capacitação de programas e organizações responsáveis pela produção e manutenção de documentos arquivísticos digitais.

---

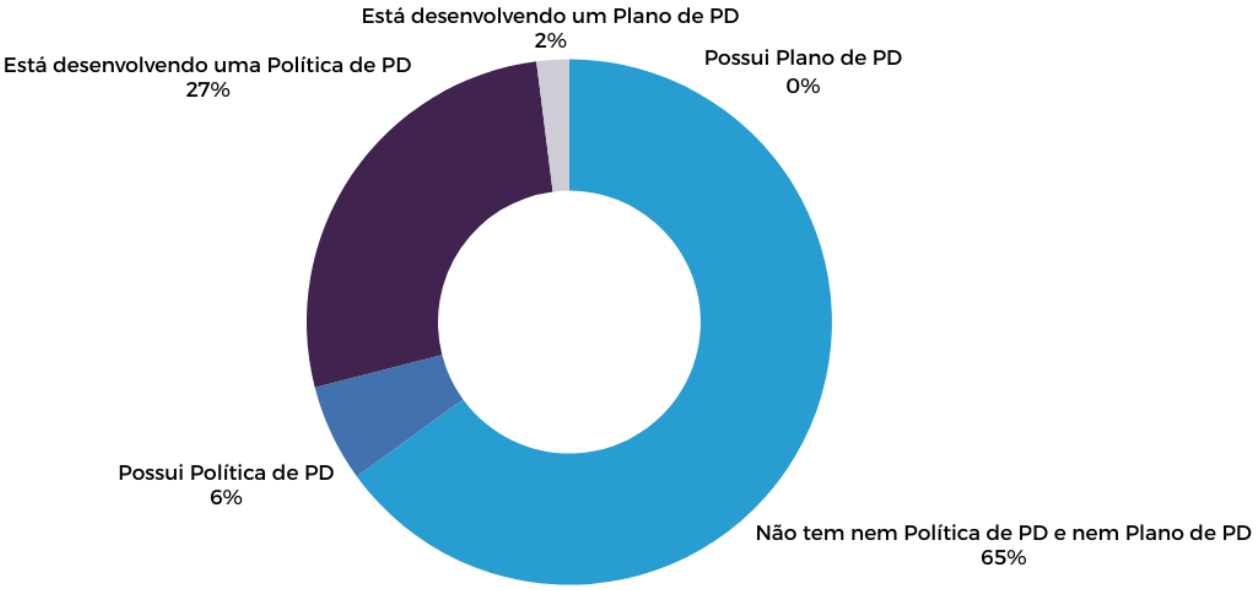
[28] A Lei n. 8.159 dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências.

[29] Diretrizes para a Preservação do Patrimônio Digital, em tradução livre.

[30] Mais informações sobre o projeto em: <https://interpares.org>

No entanto, em um estudo de 2014, "Políticas de preservação digital no Brasil: características e implementações" foi destacada a escassez de publicações de políticas de preservação digital em órgãos públicos no Brasil, com apenas dois exemplos significativos: a Política de Preservação Digital da Câmara dos Deputados e a Política de Preservação Digital da Universidade Estadual de Campinas (JÚNIOR; MOTA, 2014). Outro estudo de 2018, realizou uma pesquisa com 55 instituições públicas brasileiras e constatou que somente 6% (apenas 3) possui uma Política de preservação digital (SILVA; FLORES, 2018). A Figura 2.14 apresenta os dados percentuais obtidos nessa pesquisa.

Figura 2.14 – Preservação digital em instituições públicas federais



Fonte: adaptado de SILVA; FLORES (2018)

Desde então, o Arquivo Nacional atualizou a Política de Preservação Digital da AN Digital; em 2019, a Pinacoteca do estado de São Paulo publicou sua Política de Preservação, dividida em 4 fases para implementação e teste do *software* Archivematica em seu acervo digital; a Fundação Biblioteca Nacional (FBN)[31]

[31] A FBN é o órgão responsável pela execução da política governamental de captação, guarda, preservação e difusão da produção intelectual do Brasil.

publicou, em 2020, a Política de Preservação Digital da Biblioteca Nacional (PPDBN); a Universidade Federal do Estado do Rio de Janeiro (UNIRIO) publicou a primeira versão da sua Política de Preservação para Documentos Arquivísticos Digitais em 2022; em 2023, foi lançada a Política de Preservação Digital do Governo do Estado do Espírito Santo - PPDig@ES; e a rede SciELO atualizou sua política em 2024.

Além disso, outra análise mais recente focada nas instituições de ensino superior do Brasil mostrou que muitas delas apresentam dificuldades na implementação ou disponibilização de políticas de preservação digital. O levantamento, que incluiu as dez melhores universidades do Brasil segundo um *ranking* de 2021, encontrou uma ausência significativa de padronização e clareza nas políticas. Algumas universidades, como a Unesp, têm se destacado ao criar grupos específicos para desenvolver essas polícias (PINTO; CARNEIRO, 2024).

Na área da saúde, alguns trabalhos relevantes para a área incluem o "Plano de Preservação Digital da VideoSaúde", de BUARQUE e MACHADO (2020). O artigo descreve o processo de elaboração do Plano de Preservação Digital da VideoSaúde – Distribuidora (VSD), um serviço vinculado ao Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT), unidade técnico-científica da Fundação Oswaldo Cruz (Fiocruz). O acervo da VSD é composto majoritariamente de documentos audiovisuais e com características híbridas, tanto no que se refere à natureza dos materiais que custodia (arquivo e coleções temáticas) quanto aos suportes e meios (fitas magnéticas, discos ópticos, meios analógico e digital) que registram os seus conteúdos (BUARQUE; MACHADO, 2020).

A importância deste trabalho para a área da saúde está justamente na criação de um modelo que pode ser adaptado e replicado em outras instituições com acervos audiovisuais e híbridos, que enfrentam desafios semelhantes para preservar conteúdos valiosos de natureza científica e informativa. A pesquisa mostra como um plano bem estruturado e ajustado à realidade tecnológica e institucional de um acervo específico pode contribuir para a preservação do patrimônio digital em saúde.

O artigo supracitado foi publicado na Revista Eletrônica de Comunicação, Informação e Inovação em Saúde (RECIIS), da Fiocruz, um periódico interdisciplinar que se dedica à disseminação de conhecimento nas áreas de comunicação, informação e inovação em saúde. Em 2020, a revista lançou um dossiê temático

sobre preservação digital, com o objetivo de enriquecer o debate sobre a gestão dos acervos culturais e científicos na área da saúde (NASCIMENTO; ARAÚJO; MÁRDERO ARELLANO, 2020). No desenvolvimento desta tese, foram utilizados artigos provenientes desse dossiê, os quais contribuíram para a fundamentação teórica e discussão dos temas abordados na presente pesquisa.

No artigo "Autenticidade e preservação de Registros Eletrônicos em Saúde: proposta de modelagem da cadeia de custódia das informações orgânicas do Sistema Único de Saúde", os autores discutem a complexidade de garantir a autenticidade e preservação dos RES dentro do SUS. Meirelles e Cunha (2020) propõem uma estrutura detalhada para assegurar que esses registros mantenham sua integridade e confiabilidade ao longo de todo o ciclo de vida. A proposta inclui a gestão rigorosa da cadeia de custódia dos RES, abrangendo desde sua criação até o armazenamento e eventual descarte, sempre em conformidade com padrões rigorosos de segurança e integridade da informação.

Em complemento, o trabalho de ALMEIDA; SILVA; COSTA (2017) apresenta os resultados do projeto de digitalização e publicação *on-line* dos acervos da Coleção Memória da Enfermagem e da Nutrição da Biblioteca Setorial de Enfermagem e Nutrição (BSEN) da UNIRIO. O acervo em questão é composto de livros impressos e *e-books*, periódicos, dissertações, teses, TCC, fotografias, normas técnicas e folhetos. O artigo detalha as estratégias adotadas pela BSEN para a preservação e organização dessa coleção, assim como a utilização extensiva das notas do MARC21[32] para retratar as especificidades de cada exemplar registrado.

Além disso, também descreve as etapas de digitalização e divulgação, destacando a importância da proteção contra agentes de degradação física e digital. A equipe da BSEN optou por medidas específicas, como alocação das obras em estantes separadas e restrição do acesso físico, além da criação de uma tabela de prioridades para higienização e digitalização das obras. Critérios detalhados foram estabelecidos

---

[32] O *Machine Readable Cataloging* (MARC21) é um sistema de catalogação que foi criado para possibilitar um intercâmbio virtual de publicações (livros, periódicos e outros tipos de obras) de forma padronizada.

para a seleção de obras a serem digitalizadas, visando preservar registros históricos significativos na área da saúde. Para aumentar a visibilidade das coleções, a BSEN desenvolveu tutoriais, banners digitais e postais, além de integrar estudantes de Biblioteconomia da UNIRIO em projetos de divulgação. Um site dedicado também foi criado para promover e compartilhar o acervo da coleção memória (ALMEIDA; SILVA; COSTA, 2017).

Outro trabalho relevante na saúde é a “Proposta de modelo para a preservação e curadoria digital de objetos digitais de centros de pesquisas oncológicas". MELLO (2020) desenvolveu um modelo, baseado em adaptações de outros modelos relevantes para a área, mas adequado aos centros de pesquisas oncológicas. A pesquisa é fundamentada em um estudo de caso realizado em um centro de referência em Florianópolis, onde foram identificadas as práticas atuais e as lacunas existentes na preservação, manutenção e agregação de valor aos objetos digitais.

O panorama atual da preservação digital no Brasil revela um esforço contínuo de adaptação e crescimento, sustentada principalmente por iniciativas organizacionais e de pesquisas acadêmicas e uma crescente conscientização sobre a importância da gestão adequada dos acervos digitais. Apesar dos avanços, a realidade brasileira ainda enfrenta desafios significativos, como a escassez de políticas governamentais nos órgãos públicos e a disparidade de recursos entre as diferentes instituições.

3 Análise de valor

O objetivo principal da análise de valor é avaliar como maximizar o valor de um produto ou serviço com o menor custo possível sem sacrificar a qualidade. Originada nos anos 40 durante a Segunda Guerra Mundial, foi desenvolvida por Lawrence Miles, visando a redução de custos de fabricação de peças de aeronaves sem comprometer sua qualidade e desempenho. De acordo com a *Society of American Value Engineers* (SAVE International), "a análise de valor visa identificar e eliminar desperdícios, promovendo inovação e eficiência em diferentes áreas, desde engenharia e construção até gestão de negócios".

Rich e Holweg (2000) definem os pontos e elementos principais da definição abrangente da análise de valor:

1. A análise de valor é um processo sistemático, formal e organizado de análise e avaliação. Não é casual ou informal e é uma atividade de gerenciamento que requer planejamento, controle e coordenação.

2. A análise diz respeito à função de um produto para atender às demandas ou aplicações necessárias para um cliente. Para atender a esse requisito funcional, o processo de revisão deve incluir uma compreensão da finalidade para a qual o produto é usado.

3. O entendimento do uso de um produto implica que as especificações podem ser estabelecidas para avaliar o nível de adequação entre o produto e o valor obtido pelo cliente ou consumidor.

4. Para ser bem-sucedido, o processo de gerenciamento formal deve atender a essas especificações funcionais e os critérios de desempenho de forma consistente para agregar valor ao cliente.

5. Para gerar benefícios para a empresa, o processo de revisão formal deve resultar em um processo de melhorias no design que sirva para reduzir os custos de produção desse produto e, ao mesmo tempo, manter esse nível de valor por meio da função. (RICH; HOLWEG, 2000, tradução nossa)

O valor pode ser criado no momento da criação do produto, na sua apropriação, no seu consumo, na sua renovação ou na sua transferência. (OSTERWALDER;

PIGNEUR, 2003). O valor ao cliente, mencionado no item 4, refere-se ao valor que o cliente atribui ao produto ou serviço de acordo com sua percepção pessoal. Isso pode significar uma redução nos sacrifícios feitos pelo cliente, a presença de benefícios percebidos, ou uma combinação ponderada desses elementos, seja de forma racional ou intuitiva. Ao longo do tempo, essa percepção de valor pode se acumular e evoluir (WOODALL, 2003).

No contexto específico abordado nesta tese, os clientes incluem instituições de saúde, pesquisadores, profissionais de saúde e pacientes. O valor para esses clientes está na redução dos riscos associados à perda ou corrupção de informações críticas. Além disso, a interoperabilidade dos dados facilitará a colaboração entre diferentes entidades de saúde e a troca de informações, potencialmente resultando em avanços na área médica.

A percepção de valor para os clientes, conforme definido por Zeithaml (1988, p. 14), é "[...] a avaliação geral que o cliente faz da utilidade de um produto [, serviço ou relação] com base em suas percepções do que recebe [, os benefícios,] e do que dá em troca [, sacrifícios]". O entendimento da percepção de valor é fundamental para compreender como um produto ou serviço atendem às necessidades e expectativas dos clientes. Considerando esse conceito, no escopo desta pesquisa, o valor percebido pelos clientes se baseia na percepção dos benefícios obtidos, como a garantia de integridade e acessibilidade contínua dos dados, e dos sacrifícios realizados, como os recursos financeiros e o tempo dedicado à implementação e manutenção do plano de preservação digital. A Tabela 3.1 resume o valor percebido do modelo pretendido nesta tese.

Tabela 3.1 – Benefícios e Sacrifícios do Modelo de Preservação Digital

Produto / Domínio	**Modelo de preservação digital para base de dados na área da saúde**
Benefícios	Garantia da integridade dos dados de saúde Acessibilidade e confiabilidade das informações Interoperabilidade e colaboração facilitadas Facilidade de pesquisa e do atendimento médico Fortalecimento da posição do Brasil na saúde digital
Sacrifícios	Investimento em recursos e formação Necessidade de tempo para implementação Possível resistência à mudança organizacional Requisitos contínuos de manutenção e revisão Desafios técnicos e tecnológicos

Fonte: Elaborado pela autora, baseado em Lapierre (2000)

3.1 Proposta de Valor

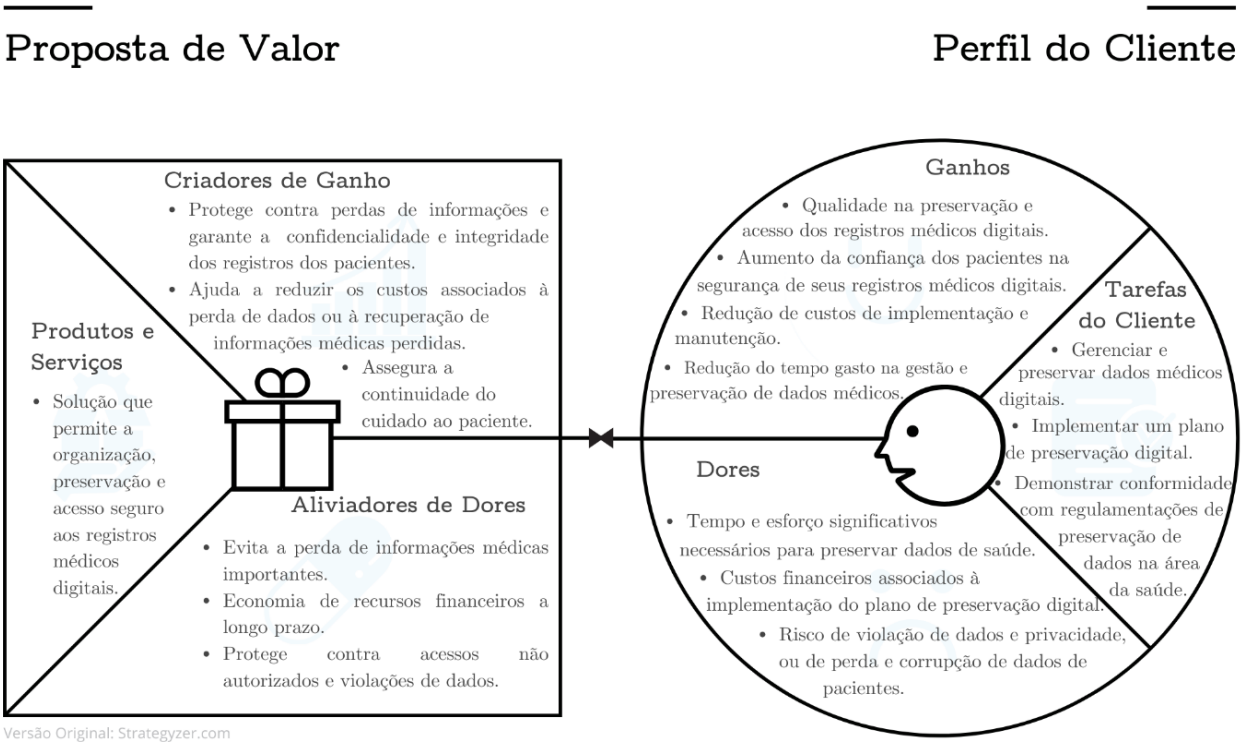
A proposta de valor, conforme definido por Osterwalder (2004), "é uma visão geral do conjunto de produtos e serviços de uma empresa que são de valor para o cliente". Ela deve ser clara, concisa e comunicar de forma convincente como a empresa se diferencia da concorrência e atende às necessidades específicas dos seus clientes. Nesse contexto, o Canvas da Proposta de Valor surge como uma ferramenta visual para garantir que um produto ou serviço esteja alinhado com os valores e necessidades do cliente, diferenciando a oferta da empresa em relação aos concorrentes e explicando por que os clientes optam por comprar dela.

No Canvas da Proposta de Valor, o segmento Perfil do Cliente descreve as Tarefas do Cliente, Ganhos e Dores. As tarefas descrevem as ações que os clientes estão tentando realizar ou completar, ou os problemas que estão tentando resolver, ou as necessidades que estão tentando satisfazer. As Dores dos clientes, descrevem as emoções negativas, custos e situações indesejadas, riscos e outras experiência ruins que o cliente pode ter vivenciado antes, durante ou após as tarefas listadas anteriormente. Por fim, os ganhos são os benefícios que os clientes esperam, desejam ou seriam surpreendidos positivamente se existissem. (OSTERWALDER et al., 2014)

No segmento Proposta de Valor, também existe três aspectos diferentes. Os Produtos e Serviços dizem respeito aos produtos e/ou serviços em que a proposta de valor está baseada. Os Aliviadores de Dores descrevem como a solução elimina ou reduz emoções negativas dos clientes, custos e situações indesejadas, riscos que os clientes vivenciam ou vivenciaram antes, durante ou após de um trabalho. Finalmente, os Criadores de Ganhos descrevem como os produtos e/ou serviços criam benefícios que os clientes desejam, esperam ou que seriam surpreendidos se existissem. (OSTERWALDER et al., 2014)

A proposta de valor do Modelo de Preservação Digital é: Garantir o gerenciamento de dados de saúde com segurança, confiabilidade e acessibilidade a longo prazo, impulsionando a pesquisa médica, a tomada de decisões clínicas e a qualidade do atendimento ao paciente. Para descrever de forma clara os recursos da proposta de valor e o segmento específico dos clientes, o Canvas da Proposta de Valor foi elaborado e pode ser visto na Figura 3.1.

Figura 3.1 – Canvas da Proposta de Valor



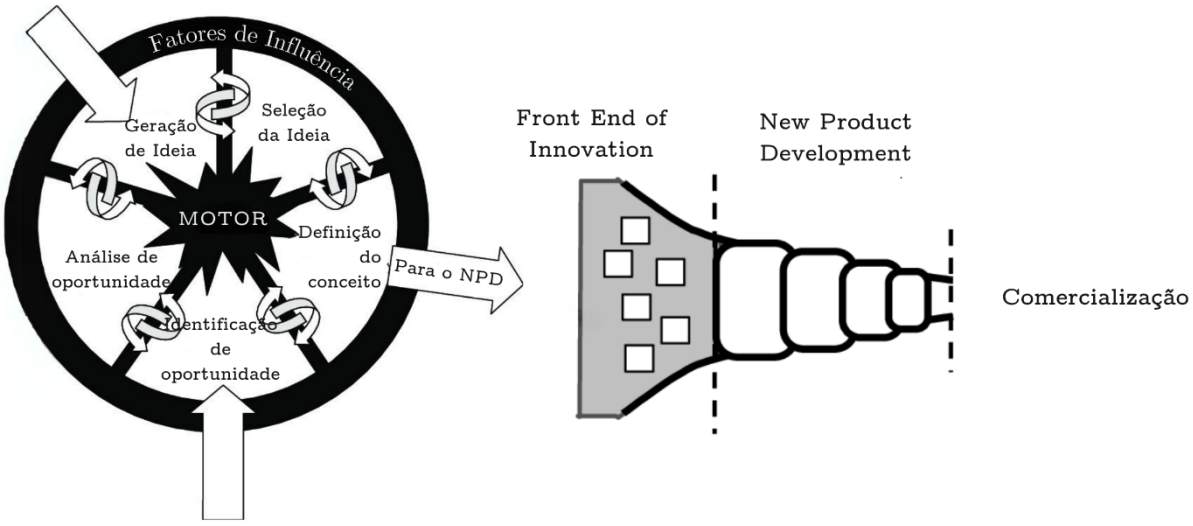
Fonte: elaborado pela autora

A partir disso, a análise de valor dessa tese utiliza o modelo *New Concept Development* (NCD), que oferece um conjunto de ferramentas para identificar problemas, definir conceitos, validá-los e apresentá-los de forma eficaz; e o Modelo Canvas do Negócio, uma extensão ao Canvas da Proposta de Valor.

3.2 Modelo New Concept Development

Segundo Koen *et al.* (2001), todo o processo de inovação pode ser dividido em três partes: *Front End of Innovation* (FEI)[33], *New Product Development* (NPD) e as fases de comercialização.

Figura 3.2 – Modelo New Concept Development e o processo de inovação



Fonte: Traduzido e adaptado de Koen et al. (2001)

O FEI é definido por Koen *et al.* (2001) como as atividades que precedem o processo formal e bem estruturado do NPD. O processo do modelo inicia-se pela geração de ideia ou identificação de oportunidade, e a forma circular do NCD, como ilustrado na Figura 3.2, indica que é esperado que as ideias fluam e se interajam entre os cinco

---

[33] Também encontrado na literatura como *Fuzzy Front End* em referência à natureza incerta e nebulosa da fase inicial do processo de inovação.

elementos. Em contrapartida, a parte do NPD é ilustrada como uma série de passos sequenciais, bem estruturados e ordenados cronologicamente.

O modelo NCD consiste em três partes principais: 1. A área interna que define os cinco elementos chave que compõem o FEI; 2. O motor que impulsiona os cinco elementos do *front-end* e é alimentada pela liderança e cultura organizacional; e 3. Os fatores de influência que consistem nas capacidades organizacionais, na estratégia do negócio, no mundo exterior (*i.e.,* canais de distribuição, clientes e concorrência) e na maturidade das tecnologias utilizadas. Esses são os mesmos fatores que afetam todo o processo de inovação, incluindo FEI, NPD e comercialização, conforme esquematizado na Figura 3.2.

No contexto desta pesquisa, serão explorados os cinco elementos do modelo NCD: Identificação de Oportunidade, Análise de Oportunidade, Geração de Ideia, Seleção de Ideia e Definição de Conceito.

IDENTIFICAÇÃO DE OPORTUNIDADE

Nesse elemento, a organização, de forma planejada ou por necessidade, identifica as oportunidades que a empresa pode querer explorar, visando alocar recursos para novas áreas de crescimento de mercado ou eficácia e eficiência operacional. Esse elemento é geralmente impulsionado pelos objetivos do negócio, podendo ser uma resposta imediata a uma ameaça competitiva, uma possibilidade de obter vantagem competitiva, ou uma maneira de reduzir ou simplificar custos operacionais (KOEN *et al.*, 2001).

Neste estágio, foi identificada a oportunidade de abordar os desafios específicos relacionados à crescente quantidade de registros digitais na área da saúde e à importância da preservação da informação médica de pacientes. A falta de um plano estruturado para lidar com o volume crescente de RES e a relevância da informação médica para diversos fins, como pesquisa, auditoria e acompanhamento do histórico do paciente, representa uma lacuna significativa na área.

ANÁLISE DE OPORTUNIDADE

Nesse estágio, informações adicionais são necessárias para traduzir a Identificação de Oportunidades em oportunidades de negócios e tecnologias específicas. Essa etapa pode demandar grupos focais, estudos de mercado e experimentos científicos, dependendo da atratividade da oportunidade, do tamanho do esforço futuro de

desenvolvimento, da adequação à estratégia e cultura do negócio e da tolerância ao risco dos tomadores de decisão (KOEN *et al.*, 2001).

A fim de explorar os aspectos internos e externos que influenciam a oportunidade e avaliar a viabilidade do produto e os desafios e oportunidades que serão enfrentados, uma Análise SWOT foi aplicada à pesquisa (Figura 3.3). Nesta análise é possível identificar as Forças (*Strengths*), Fraquezas (*Weaknesses*), Oportunidades (*Opportunities*), e Ameaças (*Threats*) relacionadas à oportunidade em questão.

Figura 3.3 – Análise SWOT

Fonte: elaborado pela autora

O primeiro aspecto a ser analisado é o ambiente em que o projeto se insere. O modelo de curadoria digital será implementado no contexto da área de saúde, o que possibilita a integração com sistemas existentes e a utilização de um amplo conjunto de dados médico. Suas forças residem na crescente necessidade de preservação de dados médicos e nos benefícios da curadoria digital, que incluem organização e proteção contra perda de dados. No entanto, as fraquezas incluem a complexidade tecnológica, a falta de familiaridade dos profissionais da saúde com a área e a necessidade de investimento em infraestrutura. As oportunidades são representadas pela crescente demanda por soluções de preservação digital, possíveis parcerias estratégicas e o aprimoramento das políticas de preservação de dados. Por outro

lado, as ameaças incluem a competição no mercado, mudanças nas tecnologias de armazenamento de dados, falta de recursos financeiros e humanos, e riscos de segurança cibernética.

GERAÇÃO DE IDEIAS

A Geração de Ideias é o estágio em que a oportunidade se transforma em uma ideia concreta, passando por um processo evolutivo de nascimento, desenvolvimento e amadurecimento. Uma nova ideia também pode surgir fora dos limites de qualquer processo formal, como um experimento que deu errado, uma oferta de um fornecedor de um novo material ou um pedido incomum de um usuário. O resultado deste estágio geralmente é uma descrição mais completa da ideia percebida ou conceito de produto (KOEN *et al.*, 2001).

Este estágio envolveu a geração de várias ideias e abordagens para a definição dos objetivos e entregas desta pesquisa. Através de discussões, foi explorado pela equipa o nível ideal de curadoria, a integração de metadados e a independência do modelo em relação a aplicações específicas. Adicionalmente, questionamentos sobre o escopo do projeto, a definição do modelo e o público-alvo fomentaram uma reflexão sobre a localização da proposta dentro do framework DAMA-DMBOK. Conforme a pesquisa se aprofundava, surgiram questões relevantes quanto à especificidade dos dados relativos à saúde de idosos. Uma análise revelou que essa particularidade não justificava a restrição do modelo. Optou-se, portanto, por uma abordagem mais abrangente, tendo como objetivo a criação de uma consultoria personalizada para uma instituição específica, considerando sua infraestrutura, recursos técnicos e financeiros. O modelo resultante, embora direcionado a uma instituição inicial, será desenvolvido de forma que possa ser adaptado e utilizado por outras organizações do setor de saúde.

Para o desenvolvimento do modelo de preservação digital, foram consideradas as três abordagens principais:

1. *Software* de Preservação

Desenvolvimento de uma aplicação ou conjunto de ferramentas digitais especificamente voltadas para a preservação de dados. Esse *software* incluiria funcionalidades como migração de formatos, verificação de integridade, armazenamento seguro e recuperação de dados.

2. Plano de Preservação

Elaboração de um plano detalhado que define estratégias, processos e práticas recomendadas para garantir a preservação digital a longo prazo. Um plano de preservação inclui as abordagens técnicas e organizacionais para a conservação de dados, uma descrição detalhada dos processos envolvidos, o conjunto de práticas para assegurar que os dados sejam armazenados e mantidos de forma segura e eficiente, a definição dos recursos técnicos, humanos e financeiros necessários para a implementação do plano, as diretrizes para garantir que os dados sejam acessíveis a pessoas autorizadas e protegidos contra acesso não autorizado, entre outros.

3. Política de Curadoria

Desenvolvimento de diretrizes e normas para a curadoria digital, abrangendo a aquisição, armazenamento, manutenção e disponibilização de dados, ou seja, durante todo o ciclo de vida de recursos digitais com o objetivo de alcançar a preservação a longo prazo. A política de curadoria deve abranger todas as atividades na gestão de dados, que pode iniciar ainda antes da criação da informação, passar por ações cíclicas de avaliação e transformação, podendo chegar à eliminação, se assim determinada por uma avaliação. (ROCHA; PIRES, 2020)

SELEÇÃO DE IDEIAS

A seguir, na Seleção de Ideias, em muitos negócios, há tantas ideias de produtos/processos que a atividade crítica é escolher quais ideias seguir para alcançar o maior valor de negócio possível (KOEN *et al.*, 2001). Para auxiliar nesse processo de tomada de decisão, foi definido o uso do *Analytic Hierarchy Process* (AHP) como a ferramenta principal.

Para tomar uma decisão de maneira organizada e gerar prioridades, precisamos decompor a decisão nos seguintes passos:

1. Definir o problema e determinar o tipo de conhecimento desejado;
2. Estruturar a hierarquia da decisão começando do topo com o objetivo da decisão, depois os objetivos de uma perspectiva ampla, passando pelos níveis intermediários (critérios dos quais os elementos subsequentes dependem) até o nível mais baixo (geralmente um conjunto de alternativas);

3. Construir um conjunto de matrizes de comparação pareada. Cada elemento em um nível superior é usado para comparar os elementos no nível imediatamente inferior em relação a ele;
4. Usar as prioridades obtidas nas comparações para pesar as prioridades no nível imediatamente inferior. Fazer isso para cada elemento. Em seguida, para cada elemento no nível inferior, adicionar seus valores ponderados e obter sua prioridade total ou global. (SAATY, 2008)

Os critérios definidos para comparação nessa pesquisa são: Custo, Viabilidade Técnica, Eficácia, Sustentabilidade e Usabilidade. A Tabela 3.2 apresenta uma descrição geral de cada critério.

Tabela 3.2 – Critérios de comparação

Critério	**Descrição**
Custo	Envolve o custo para a instituição implementar e manter a solução
Viabilidade Técnica	Avaliação da infraestrutura técnica requerida, o grau de dificuldade técnica para implementação e a compatibilidade com sistemas e formatos existentes
Eficácia	Relacionada com a efetividade, flexibilidade e escalabilidade da solução
Sustentabilidade	Necessidade de recursos humanos especializados para a implementação e manutenção
Usabilidade	Quão intuitiva e fácil de acessar é a solução para os usuários

Para fazer comparações, precisamos de uma escala numérica que indique quantas vezes um elemento é mais importante em relação a outro elemento, considerando o critério ou propriedade sob os quais eles são comparados. A Tabela 3.3 apresenta essa escala. A escala de comparação de pares usada no AHP varia de 1 a 9, onde 1 indica igual importância e 9 indica importância extrema de um elemento sobre outro. Valores intermediários (2, 4, 6, 8) são usados para representarem importâncias moderadas entre esses extremos.

Tabela 3.3 – A Escala Fundamental

Nível de Importância	**Definição**	**Explicação**
1	Igual Importância	As duas atividades contribuem igualmente para o objetivo
3	Importância moderada	A experiência e o julgamento favorecem ligeiramente uma atividade em relação a outra
5	Forte importância	A experiência e o julgamento favorecem fortemente uma atividade em relação a outra
7	Muito forte importância	A experiência e o julgamento favorecem muito fortemente uma atividade em relação a outra
9	Extrema importância	A evidência que favorece uma atividade em relação a outra com o mais alto grau de certeza
2, 4, 6, 8	Valores intermediários	Quando se procura uma condição de compromisso entre duas definições.

Fonte: SAATY (1991)

A segunda fase consiste em estabelecer prioridades entre os elementos por meio de uma matriz de comparação a partir dos valores determinados na Escala Fundamental. A seguir, na Tabela 3.4, é apresentada a matriz preenchida de acordo com o método AHP.

Tabela 3.4 – Matriz de comparação dos critérios

Critérios	**Viabilidade Técnica**	**Eficácia**	**Sustentabilidade**	**Custo**	**Usabilidade**
Viabilidade Técnica	1	3	5	7	9
Eficácia	1/3	1	3	5	7
Sustentabilidade	1/5	1/3	1	3	5
Custo	1/7	1/5	1/3	1	3
Usabilidade	1/9	1/7	1/5	1/3	1

No caso dessa análise, observa-se que a Viabilidade Técnica é 3 vezes mais dominante que Eficácia, 5 vezes mais que Sustentabilidade, 7 vezes mais dominante que Custo e 9 vezes mais dominante que Usabilidade, tornando-se o critério mais relevante para a escolha da abordagem desta pesquisa. No entanto, ainda é necessário normalizar esses valores e igualar todos os critérios a uma mesma unidade. Para isso, na fase 3, cada valor da matriz é dividido pelo total da sua respectiva coluna e pode ser visto na Tabela 3.5. Todos os cálculos foram realizados utilizando a linguagem de programação *Python*, o código pode ser acessado no *Google Colaboratory*.

Tabela 3.5 – Matriz de comparação normalizada dos critérios

Critérios	**Viabilidade Técnica**	**Eficácia**	**Sustentabilidade**	**Custo**	**Usabilidade**
Viabilidade Técnica	315/563	315/491	75/143	3/7	9/25
Eficácia	105/563	105/491	45/143	15/49	7/25
Sustentabilidade	63/563	35/491	15/143	9/49	1/5
Custo	45/563	21/491	5/143	3/49	3/25
Usabilidade	35/563	15/491	3/143	1/49	1/25
Soma	563/315	491/105	143/15	49/3	25

Para identificar a ordem de importância de cada critério, é calculada a média aritmética dos valores de cada linha da matriz normalizada obtida na Tabela 3.5. Os valores obtidos desse cálculo estão apresentados na Tabela 3.6.

Tabela 3.6 – Matriz de comparação com a Prioridade relativa

Critérios	**Viabilidade Técnica**	**Eficácia**	**Sustentabilidade**	**Custo**	**Usabilidade**	**Prioridade Relativa**
Viabilidade Técnica	315/563	630/491	75/143	3/7	9/25	0.50
Eficácia	105/563	105/491	45/143	15/49	7/25	0.26
Sustentabilidade	63/563	35/491	15/143	9/49	1/5	0.13
Custo	45/563	21/491	5/143	3/49	3/25	0.07
Usabilidade	35/563	15/491	3/143	1/49	1/25	0.03

Seguindo o grau de importância, do mais relevante ao menos relevante, temos a seguinte ordem: 1. Viabilidade Técnica; 2. Eficácia; 3. Sustentabilidade; 4. Custo; e 5. Usabilidade. Para determinar a alternativa mais adequada, é necessário realizar a comparação de pares entre as abordagens e considerando cada critério. O mesmo procedimento foi realizado para cada tabela abaixo.

Tabela 3.7 – Matriz de comparação das abordagens para o critério Custo

Alternativa	**Software de Preservação**	**Plano de Preservação**	**Política de Curadoria**
Software de Preservação	1	1/5	1/3
Plano de Preservação	5	1	3
Política de Curadoria	3	1/3	1

Tabela 3.8 – Matriz de comparação das abordagens para o critério Viabilidade Técnica

Alternativa	**Software de Preservação**	**Plano de Preservação**	**Política de Curadoria**
Software de Preservação	1	3	5
Plano de Preservação	1/3	1	3
Política de Curadoria	1/5	1/3	1

Tabela 3.9 – Matriz de comparação das abordagens para o critério Eficácia

Alternativa	**Software de Preservação**	**Plano de Preservação**	**Política de Curadoria**
Software de Preservação	1	3	5
Plano de Preservação	1/3	1	3
Política de Curadoria	1/5	1/3	1

Tabela 3.10 – Matriz de comparação das abordagens para o critério Sustentabilidade

Alternativa	**Software de Preservação**	**Plano de Preservação**	**Política de Curadoria**
Software de Preservação	1	5	7
Plano de Preservação	1/5	1	3
Política de Curadoria	1/7	1/3	1

Tabela 3.11 – Matriz de comparação das abordagens para o critério Usabilidade

Alternativa	**Software de Preservação**	**Plano de Preservação**	**Política de Curadoria**
Software de Preservação	1	1/3	1/5
Plano de Preservação	3	1	1/3
Política de Curadoria	5	3	1

Ao normalizar cada uma das matrizes de comparação de alternativas, calcular os pesos das alternativas para cada critério e combinar os pesos dos critérios com os pesos das alternativas, a pontuação final para cada alternativa teve o seguinte resultado:

Tabela 3.12 – Pontuação final das abordagens

Alternativa	Pontuação
Software de Preservação	0.36
Plano de Preservação	0.44
Política de Curadoria	0.20

Com base na aplicação do método APH, a melhor abordagem para o desenvolvimento de um modelo de preservação digital é o **Plano de Preservação**.

Dessa forma, foi definido para esta pesquisa a elaboração de um plano de preservação digital que esteja em conformidade com os padrões e legislações vigentes, e que garanta a segurança, integridade e o acesso contínuo aos dados da área de saúde.

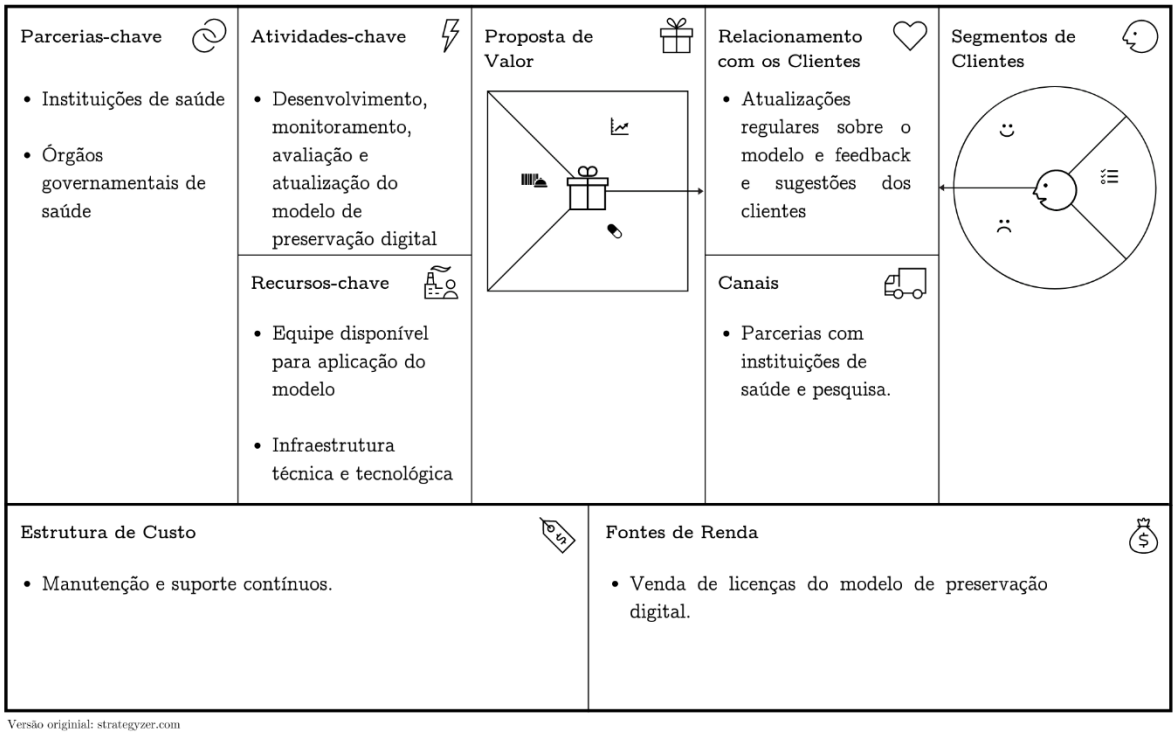
DEFINIÇÃO DE CONCEITO E TECNOLOGIA

O Desenvolvimento de Conceito e Tecnologia representa o último elemento do modelo e envolve a elaboração de um caso de negócio com base em estimativas de potencial de mercado, necessidades do cliente, requisitos de investimento, avaliações da concorrência, tecnologias associadas e riscos do projeto como um todo. Para esse estágio, o Canvas do Modelo de Negócios (ou *Business Model Canvas*) foi utilizado para estruturar e visualizar o modelo de negócio do projeto e será apresentado na próxima seção.

3.3 Canvas do Modelo de Negócios

Um Modelo de Negócios descreve a lógica de criação, entrega e captura de valor por parte de uma organização. OSTERWALDER; PIGNEUR; CLARK (2010) acreditam que um modelo de negócios pode ser mais bem descrito com nove componentes básicos que cobrem as quatro principais áreas de um negócio: clientes, oferta, infraestrutura e viabilidade financeira. Esses nove componentes formam a base para uma ferramenta útil: o Canvas (ou Quadro) do Modelo de Negócios. A Figura 3.4 exemplifica a aplicação deste quadro à proposta desta tese.

Figura 3.4 – Canvas do Modelo de Negócios



Fonte: elaborado pela autora

Os componentes **Segmentos de Clientes** e **Proposta de Valor** foram definidos na seção 3, na Figura 3.1 com o Canvas da Proposta de Valor e que complementa a elaboração do Canvas do Modelo de Negócios. O componente **Canais** descreve como uma empresa se comunica e alcança seus Segmentos de Clientes para entregar uma Proposta de Valor. Os canais definidos nessa proposta estão classificados como canais de parceria, uma vez que é essencial parcerias com instituições de saúde e pesquisa para desenvolver e aplicar o modelo desenvolvido.

O **Relacionamento com Clientes** descreve os tipos de relação que uma empresa estabelece com Segmentos de Clientes específicos. O relacionamento adotado baseia-se na assistência pessoal dedicada e cocriação, com atualizações regulares sobre o modelo elaborado e aceitação de *feedback* e sugestões do cliente. As **Fontes de Receita** serão transações de renda resultantes de pagamento único com as vendas de licença do modelo de preservação digital. Os **Recursos-chave** abrangem os recursos humanos, financeiros e físicos que serão necessários para implementação do modelo, são eles uma equipe disponível para aplicação e teste do modelo e infraestrutura técnica e tecnológica.

As **Atividades-chave** são as ações mais importantes que deverão ser executadas para operar com sucesso, no contexto dessa pesquisa as ações de desenvolvimento, monitoramento, avaliação e atualização do modelo de preservação digital são as ações principais do projeto. As **Parcerias-chave** são as instituições de saúde e órgãos governamentais de saúde. Por fim, a **Estrutura de Custo** está principalmente ligada à manutenção e suporte contínuos à instituição.

4 Experimentação

A experimentação consiste em obter resultados, enquanto a avaliação analisa esses resultados, concluindo e possivelmente redirecionando a implementação da solução. Essa tese tem como objetivo a implementação de um modelo de preservação digital, adaptado às características específicas ao ambiente institucional de saúde. O modelo deverá garantir os seguintes objetivos:

1. Assegurar a preservação e o acesso contínuo aos objetos digitais;
2. Garantir a privacidade dos dados, restringindo o acesso apenas a usuários credenciados;
3. Verificar a autenticidade dos materiais preservados;
4. Preservar a mídia física contra danos e deterioração por meio de controles ambientais adequados;
5. Reverter os danos, quando possível;
6. Adaptar o formato ou padrões (de metadados ou de interoperabilidade) dos objetos digitais para preservar seu conteúdo lógico e intelectual, quando necessário.

O conceito de modelo na proposta desta tese se alinha à função de representar e explicar realidades complexas no contexto de saúde. Segundo JAPIASSU; MARCONDES (2008), um modelo atua como um paradigma ou uma construção teórica que auxilia na compreensão e análise ou avaliação de uma realidade concreta, como por exemplo os desafios relacionados à preservação de dados de saúde. Nesse contexto, o modelo de preservação digital proposto é uma estrutura teórica que reflete os processos, técnicas e estratégias necessárias para manter a integridade, acessibilidade e usabilidade desses dados ao longo do tempo.

Esse modelo deve ser visto também como uma ferramenta para descrever de forma específica os dados em saúde. Essa abordagem implica a construção de um modelo de preservação digital capaz de lidar com os tipos de dados, interoperabilidade de sistemas, questões de segurança e privacidade, e a capacidade de migração de informações. Dessa forma, o modelo proposto não apenas organiza o conhecimento

existente, mas também orienta a implementação de práticas para a preservação digital no setor da saúde.

Neste sentido, o modelo elaborado nesta tese será um plano de preservação digital, que se diferencia de políticas de alto nível, que são geralmente formuladas em nível institucional e regulam variáveis fundamentais e estratégias abrangentes. Por outro lado, um plano de preservação é mais específico e concreto, pois define um plano de ação detalhado para preservar um conjunto particular de objetos com um determinado objetivo. (BECKER *et al.* 2009)

BECKER *et al.* (2009) define que um plano de preservação deve conter os seguintes elementos: Identificação; Estado e Gatilhos; Descrição do ambiente institucional; Descrição da coleção; Requisitos para preservação; Evidência de decisão para uma estratégia de preservação; Custos; Funções e Responsabilidades; e por fim, Plano de ação de preservação. Com base nesses elementos e nas recomendações gerais de boas práticas da área expostas na *Digital Preservation Coalition*, o seguinte plano foi delimitado para essa pesquisa:

- Avaliação institucional e Planejamento
    - Descrição da coleção atual

      Uma coleção é o conjunto de objetos digitais ou registros para os quais um plano de preservação é criado. Para montar uma infraestrutura tecnológica a instituição precisa conhecer os seus ativos digitais: a quantidade de objetos digitais e o tamanho dos arquivos; se estes objetos são simples ou complexos; o valor que cada objeto representa para a instituição; e qual o tipo de controle deve ser exercido sobre esse objeto, por exemplo, se há um tempo mínimo de permanência no acervo. (GALINDO; LA FUENTE, 2015).

      Essa etapa requer uma descrição detalhada da coleção, como também especificado em "Descrição da coleção" de BECKER *et al.* (2009). Isso inclui informações sobre os tipos de materiais, formatos de dados, volumes, condições de conservação e outras características relevantes da coleção que serão preservadas digitalmente. Além disso, também inclui objetos amostrais que são representativos para a coleção e, portanto, podem ser usados para o processo de avaliação. Isso deve

incluir os objetos reais e uma descrição de suas propriedades bem compreendidas, bem como seu ambiente técnico original.

- Avaliação da maturidade de Preservação digital

  Essa etapa visa identificar áreas de melhoria e assegurar que as melhores práticas e técnicas estão sendo aplicadas no contexto da instituição escolhida. O foco principal será a identificação de estratégias e tecnologias de preservação digital utilizadas na instituição. Adicionalmente, será verificado se as práticas da instituição estão alinhadas com recomendações e padrões reconhecidos.

  A avaliação incluirá uma revisão dos metadados gerenciados pela instituição e verificação do fluxo de trabalho definido, desde a ingestão até o armazenamento e acesso aos dados digitais. Com base nisso, será determinado um nível de maturidade de preservação digital na instituição.

- Definir objetivos e escopo

  É importante definir claramente os objetivos e o escopo do plano de preservação digital. Isso inclui identificar o propósito da preservação, os critérios de seleção de itens a serem preservados e uma especificação das propriedades significativas dos objetos, e os resultados esperados.

- Gestão de Riscos

  - Avaliar os riscos

    Essa etapa está diretamente relacionada com "Evidência de decisão para uma estratégia de preservação" pois envolve a análise e decisão sobre as estratégias de preservação. Isso inclui identificar ameaças potenciais, descrever possíveis perdas de informação, e delinear estratégias para mitigar esses riscos, fundamentando a escolha das estratégias de preservação mais adequadas.

  - Desenvolver uma estratégia de preservação

Com base na avaliação de riscos, uma estratégia que inclua medidas preventivas e ações a serem tomadas quando os riscos se materializarem será criada.

- Plano de ação da Preservação digital
    - Escolher ferramentas e padrões adequados

      As ferramentas serão avaliadas em termos de funcionalidade, interoperabilidade, comunidade de suporte, e sustentabilidade a longo prazo. Os padrões abertos são preferidos para garantir a interoperabilidade e a capacidade de migração futura. Essa etapa pode incluir a definição de ferramentas para recuperação de dados, *software* para conversão de formatos de arquivos, padrões de metadados e interoperabilidade, e solução para armazenamento.

    - Desenvolver procedimentos

      Essa etapa especifica as ações que serão aplicadas aos objetos digitais. Inclui as atividades diárias de preservação digital, como ingestão de novos materiais digitais, realização de verificações regulares de integridade e gerenciamento de metadados. As estratégias técnicas e procedimentos de preservação também serão definidas nessa etapa.

- Monitoramento e Revisão
    - Monitoramento regular

      Essa etapa requer o estabelecimento de uma estratégia para monitorar continuamente a eficácia das ações de preservação e para identificar quando mudanças ou intervenções são necessárias.

    - Revisão e atualização periódica

      Envolve a identificação de indicadores que disparam a necessidade de revisão e atualização do plano. Os indicadores poderão incluir taxas de sucesso na migração de formatos, confiabilidade das verificações de integridade e eficiência de armazenamento.

5 Modelo de Preservação Digital

A preservação de dados digitais na área da saúde exige soluções que considere a diversidade e sensibilidade dos dados envolvidos. Neste capítulo, será apresentado o modelo de preservação digital desenvolvido a partir das diretrizes conceituais e práticas discutidas nos capítulos anteriores. O objetivo é fornecer uma estrutura clara e adaptável que assegure a preservação de bases de dados de saúde. O modelo é caracterizado como um Plano de Preservação Digital, documento necessário para orientar as ações de preservação digital (FERREIRA, 2006; BECKER *et. al.,* 2009; MÁRDERO ARELLANO, 2012; BROWN, 2013; SANTOS; FLORES, 2015a; OWENS, 2018; GRACIO, 2019).

A seguir, serão discutidos os componentes fundamentais do modelo de preservação digital, incluindo as etapas de avaliação institucional, planejamento, gestão de riscos, escolha de ferramentas e padrões, e monitoramento contínuo. Cada um desses elementos será apresentado em detalhes, com base nas melhores práticas da área, para orientar a implementação de uma estratégia eficiente de preservação digital para as instituições de saúde.

5.1 Avaliação Institucional

A avaliação institucional é o processo inicial que permite à organização compreender seu estado atual em relação à gestão e preservação de seus ativos digitais. Para GRÁCIO (2012), o primeiro passo é avaliar e levantar as informações e os objetos digitais que devem ser preservados, de acordo com as necessidades e objetivos da instituição. O objetivo é identificar os pontos fortes, fraquezas, oportunidades e ameaças que possam impactar o plano de preservação digital a ser implementado. No contexto de um ambiente OAIS, essa etapa permite que a entidade de Gestor possa tomar decisões fundamentadas sobre a preservação digital.

A Unidade de Saúde São Lucas, uma unidade fictícia de atenção especializada de média complexidade localizado no Brasil, enfrenta os desafios comuns do setor público de saúde nacional no que se refere à preservação digital. Com capacidade

para atender mais de 350 pacientes diariamente, a rotina hospitalar envolve a geração e gestão de um volume significativo de dados clínicos, administrativos e operacionais, o que inclui desde prontuários eletrônicos até exames de imagem e dados financeiros.

A preservação digital, porém, não é uma prioridade ativa na instituição, devido principalmente a falta de profissionais especializados e os recursos financeiros limitados que afetam diretamente a implementação de estratégias de preservação de longo prazo. A gestão da instituição é operada em nível municipal pela Secretaria Municipal de Saúde, e seria de sua responsabilidade a implementação do plano de preservação digital. No entanto, órgãos governamentais, como a Secretaria Estadual de Saúde e o Ministério da Saúde podem oferecer um papel de suporte, especialmente em termos de regulamentação e financiamento.

5.1.1 Descrição da coleção atual

A Unidade de Saúde São Lucas lida com uma vasta coleção de objetos digitais e registros eletrônicos. A coleta inicial dos dados revela que esses ativos incluem objetos nato-digitais e digitalizados, arquivos de imagens diagnósticas, arquivos multimídia e documentos administrativos e financeiros.

Para fornecer uma descrição mais completa da coleção atual do hospital, adotou-se um formato inspirado no modelo de inventário de preservação digital (CHIN, 2017) do *The Canadian Heritage Information Network* (CHIN), desenvolvido para instituições de patrimônio cultural. Esse formato começa com a descrição dos principais elementos dos objetos digitais, apresentado na Tabela 5.1, que é dedicada a identificar e resumir todos os grupos desses objetos mantidos pela instituição. Conforme definido pelo CHIN (2017), um grupo de ativos digitais refere-se a um conjunto de ativos que compartilham o mesmo propósito, formato de arquivo ou *software* utilizado. Posteriormente, deve ser apresentado um detalhamento das informações específicas de cada objeto.

Tabela 5.1 – Resumo dos objetos digitais da Unidade de Saúde São Lucas

Grupos de objetos digitais	**Descrição**	**Número aproximado de objetos digitais**	**Número aproximado de espaço de armazenamento**	**Número de cópias mínimas**
Objetos Nato-digitais e Digitalizados relativos ao paciente	Prontuários eletrônicos de pacientes (PEC), exames laboratoriais, documentos de alta, relatórios de procedimentos médicos, e documentos digitalizados, que incluem, principalmente, as fichas de Coleta de Dados Simplificada.	150 a 200 prontuários diários	Aproximadamente 2TB anuais	Uma cópia armazenada localmente e na base nacional de dados
Imagens Clínicas e Lâminas Digitalizadas	Exames de imagem, como radiografias, tomografias e ressonâncias magnéticas, gerenciados pelo sistema PACS, além da digitalização de lâminas de exames	Cerca de 150 exames por dia	5 a 10TB por ano	Uma cópia armazenada localmente, sem backup externo formal

	histopatológicos, citopatológicos e outros tipos de lâminas utilizadas para diagnósticos laboratoriais			
Multimídia	Vídeos de tele consulta e gravações de exames	Cerca de 50 vídeos diários	1TB ao ano	
Dados Administrativos e Financeiros	Relatórios financeiros, planilhas de custos, faturamento e registros de pacientes e agendamentos, gerenciados pelo Sistema Nacional de Regulação (SISREG)	Relatórios e registros gerados diariamente	Cerca de 500GB por ano	

Fonte: elaborado pela autora

O detalhamento nas tabelas abaixo aborda outros aspectos importantes para a preservação digital associados à coleção digital. A "facilidade de substituição", por exemplo, avalia a complexidade de repor os dados em caso de perda, o que depende da natureza e do contexto em que esses ativos foram criados. Para alguns grupos de ativos, como documentos que podem ser recriados a partir de registros ou fontes existentes, a substituição pode ser relativamente fácil. No entanto, existem situações em que a reposição pode envolver redigitalização, o que frequentemente implica custos elevados, além de possíveis danos aos materiais originais, como no caso de

manuscritos históricos ou registros médicos. Em outras situações, determinados ativos digitais não podem ser substituídos de forma alguma.

O “impacto da perda” pode variar de consequências menores até graves prejuízos institucionais, dependendo do tipo de objeto digital perdido. Em casos em que a perda de dados resulta em impacto grave, como documentos clínicos ou informações que afetam a reputação da instituição, os danos podem incluir questões legais e comprometer a continuidade dos serviços. Os "anos estimados para preservação" orientam o planejamento de longo prazo, ajudando a estabelecer metas de preservação adequadas, principalmente devido a requisitos legais; enquanto a "frequência de acesso" descreve com que frequência os ativos digitais são acessados, o que ajuda a definir quais objetos exigem maior disponibilidade.

Outros aspectos, como "permissão de acesso", "formatos físicos" e "segurança" mapeiam quem pode acessar os dados, em quais condições, e quais medidas de proteção estão implementadas. “Formatos de arquivo”, “nomeação de arquivos”, “estrutura de diretórios” especificam os tipos de arquivos utilizados, o padrão de nomenclatura e sua organização. Por fim, as “medidas de preservação” descrevem as práticas já adotadas pela instituição e que colaboram para a preservação a longo prazo dos objetos digitais. O detalhamento é mostrado nas tabelas que seguem.

Tabela 5.2 – Detalhamento objetos nato-digitais e digitalizados

Aspecto	**Detalhamento**
Facilidade de substituição	**Não pode ser substituído**. Os prontuários eletrônicos de pacientes e documentos digitalizados são informações únicas que, em caso de perda, não podem ser recriados, pois envolvem dados clínicos e históricos do paciente. A substituição seria impossível ou implicaria em grandes perdas de dados médicos críticos.
Consequência / Impacto da perda	**Impacto grave.** A perda de documentos clínicos pode ter consequências legais significativas, além de prejudicar o atendimento ao paciente, pois essas

	informações são essenciais para continuidade de tratamentos e diagnósticos.
Anos estimados de preservação	**Indefinido**. Pela legislação, os prontuários devem ser preservados permanentemente.
Frequência de acesso	**Diariamente**. Esses documentos são acessados diariamente por médicos e profissionais de saúde, e em casos de emergência médica.
Permissão de acesso	**Indivíduos específicos dentro da instituição**. Apenas profissionais de saúde e administradores autorizados têm acesso a esses dados, devido às restrições de privacidade e sigilo médico.
Formatos físicos / suportes físicos	**Servidores.** Os dados são armazenados em servidores do hospital.
Formatos de arquivo / Tipo de arquivo	**.pdf, .doc/.docx, .csv, .txt., formatos proprietários do PEC** Documentos em diversos formatos textuais. O sistema PEC utiliza formatos de banco de dados específicos.
Nomeação de arquivos	**Sistema padronizado interno**. O sistema de nomeação de arquivos segue padrões do sistema PEC, gerados automaticamente. Documentos não tem um padrão de nomeação.
Estruturas de diretórios	**Dados estruturados**. Os arquivos são organizados de forma hierárquica e estruturada, com registros de pacientes e documentos clínicos associados.
Segurança	**Protegido por senha, criptografia e controle de acesso**. Os sistemas

	utilizam senha e criptografia para garantir a segurança dos dados sensíveis.
Medidas de preservação	***Backup* regular**. Realização de *backups* frequentes.

Fonte: elaborado pela autora

Tabela 5.3 – Detalhamento de arquivos de imagens diagnósticas

Aspecto	**Detalhamento**
Facilidade de substituição	**Substituição envolve re-digitalização, difícil e cara.** A reobtenção de exames seria complexa, exigindo a repetição dos exames médicos, o que pode ser prejudicial ao paciente e caro para o hospital.
Consequência / Impacto da perda	**Impacto moderado a grave**. Perda de imagens diagnósticas pode atrasar ou comprometer diagnósticos e tratamentos, além de gerar possíveis repercussões legais.
Anos estimados de preservação	**Indefinido**. Exames de imagens fazem parte do prontuário do paciente e são mantidas permanentemente.
Frequência de acesso	**Semanalmente**. Acesso regular por médicos para diagnóstico e revisão de tratamentos.
Permissão de acesso	**Profissionais de saúde e técnicos autorizados**. Apenas radiologistas, médicos e técnicos específicos podem acessar esses arquivos, devido à confidencialidade.
Formatos físicos / suportes físicos	**Armazenamento no PACS e servidores locais**. Os arquivos de imagem são armazenados no sistema PACS, utilizando servidores locais.

Formatos de arquivo / Tipo de arquivo	DICOM.
Nomeação de arquivos	**Gerado pelo PACS automaticamente**. A nomeação dos arquivos é controlada automaticamente pelo sistema PACS.
Estruturas de diretórios	**Estruturado por exames e pacientes**. O PACS organiza as imagens de acordo com o exame e o paciente, com diretórios estruturados.
Segurança	**Protegido por senha, criptografia e controle de acesso**. O sistema PACS utilizado é protegido por criptografia, senhas e acessos restritos a usuários autorizados.
Medidas de preservação	***Backup* regular**. Realização de *backups* regularmente.

Fonte: elaborado pela autora

Tabela 5.4 – Detalhamento de arquivos multimídia

Aspecto	**Detalhamento**
Facilidade de substituição	**Difícil de substituir**. Os vídeos de consultas e exames são gravações únicas, e sua substituição não é viável sem repetição do processo.
Consequência / Impacto da perda	**Impacto moderado**. A perda de gravações pode impactar revisões de casos médicos e gerar dificuldades para auditoria de procedimentos ou consultas.
Anos estimados de preservação	**1 a 5 anos**. Esses arquivos geralmente são preservados por um período menor.
Frequência de acesso	**Mensal**. Os vídeos são acessados em situações específicas, para revisão de diagnósticos ou treinamento.
Permissão de acesso	**Indivíduos específicos dentro da**

	instituição. Apenas médicos, técnicos e pessoal autorizado têm acesso aos vídeos por questões de privacidade e sigilo médico.
Formatos físicos / suportes físicos	Armazenamento em rede local e servidores.
Formatos de arquivo / Tipo de arquivo	**.mp4, .avi, .wav.** A maioria dos vídeos está armazenado nesse formato.
Nomeação de arquivos	**Não segue um padrão definido**. A nomeação de arquivos é feita de forma inconsistente, variando de acordo com o criador ou situação.
Estruturas de diretórios	**Estrutura cronológica por paciente e data**. O sistema organiza os arquivos de acordo com a data de criação e identificação do paciente.
Segurança	**Acesso restrito e criptografia**. Além de *firewall* e controle de senha, o sistema de vídeo conta com criptografia para proteger os arquivos.
Medidas de preservação	***Backup* regular**. Realização de *backups* regularmente.

Fonte: elaborado pela autora

Tabela 5.5 – Detalhamento de dados administrativos e financeiros

Aspecto	Detalhamento
Facilidade de substituição	**Pode ser substituído com algum esforço**. Muitos registros podem ser recriados a partir de sistemas financeiros e relatórios, mas a perda ainda seria prejudicial.
Consequência / Impacto da perda	**Impacto moderado**. A perda de dados administrativos pode afetar a gestão financeira do hospital, gerar atrasos e

	problemas operacionais.
Anos estimados de preservação	**Até 30 anos**. Documentos institucionais requerem tempo mínimo diferentes para cada tipo de documento.
Frequência de acesso	**Semanalmente**. A equipe administrativa acessa esses registros com frequência para operações financeiras e contabilidade.
Permissão de acesso	**Administradores e equipe financeira**. Apenas a equipe financeira e de gestão tem acesso a esses dados.
Formatos físicos / suportes físicos	**Servidores e SISREG**. Os dados são armazenados em servidores locais, integrados ao do hospital.
Formatos de arquivo / Tipo de arquivo	**.xlsx, .csv, .pdf, formatos proprietários do SISREG**. Esses arquivos estão em diferentes formatos e alguns específicos do SISREG.
Nomeação de arquivos	**Gerado pelo SISREG automaticamente**. A nomeação segue padrões automáticos definidos pelo sistema.
Estruturas de diretórios	Estrutura hierárquica de relatórios.
Segurança	**Senha e firewall**. O sistema é protegido com senha, controle de acessos e *firewall*.
Medidas de preservação	***Backup* regular**. Realização de *backups* regularmente.

Fonte: elaborado pela autora

A descrição detalhada da coleção digital da instituição serve como base para o planejamento de uma estratégia de preservação digital, e leva em conta os riscos de perda de dados, assim como as exigências operacionais e legais. Ao compreender a diversidade de formatos, a complexidade de substituição e os requisitos de acesso, é

possível avaliar a importância da coleção, definir prioridades e planejar ações de preservação de forma mais direcionada para a Unidade de Saúde São Lucas.

5.1.2 Avaliação de maturidade de preservação digital

O processo de avaliação de maturidade de preservação digital auxilia a instituição a compreender sua capacidade em preservar dados a longo prazo e a identificar áreas de melhorias. O objetivo desse processo é analisar as capacidades organizacionais, técnicas e funcionais da instituição, de forma a orientar o desenvolvimento de um plano de preservação digital. Para avaliar esses aspectos, foi utilizado principalmente o Modelo de Avaliação Rápida do DPC[34] alinhado às legislações e normas brasileiras de saúde digital.

O DPC RAM, revisado e atualizado em março de 2024, se baseia em uma variedade de modelos de maturidade existentes, entre eles estão: o Modelo de Maturidade de Adrian Brown, os Níveis de Preservação do NDSA e o Modelo de Maturidade da Capacidade de Preservação Digital (DPCMM). A escolha pelo DPC RAM como referência principal se justifica pelo fato de que, diferente de outros modelos de maturidade existentes, não tem uma área ou setor de atuação específico como alvo, nem limitam seu escopo a um subconjunto de considerações e abordagens de preservação específicas (DPC, 2024).

O modelo avalia 11 áreas de competência divididas em dois grupos que abrangem as principais áreas de preservação digital. As primeiras 6 são as "Capacidades organizacionais", que descrevem o quão apta a organização está para gerenciar as atividades de preservação digital (como recursos alocados, políticas e suporte). Os outro 5 são as "Capacidades de serviço", que descrevem os processos de preservação implementados na organização (como aquisição de dados, preservação de integridade

[34] The Digital Preservation Coalition Rapid Assessment Model (DPC RAM), em inglês.

dos *bits* e acesso). Cada uma dessas capacidades é avaliada em uma escala de 0 a 4, sendo que 0 indica uma conscientização mínima dos problemas destacados nessa área do modelo e 4 indica que a organização está trabalhando em um nível otimizado.

Tabela 5.6 – Avaliação de Maturidade de Preservação Digital

Capacidades organizacionais		
	Nível de Maturidade	**Justificativa**
A. Viabilidade Organizacional: Governança, estrutura organizacional, recursos humanos e financeiros para atividades de preservação digital.	2 – Básico	A Unidade de Saúde São Lucas parece estar ciente da necessidade de atividades de preservação digital, mas essa consciência ainda não se traduziu em ações organizacionais. Apesar da existência de algumas práticas operacionais, como *backups* e criptografia de dados, não há evidências claras de um planejamento estratégico consolidado ou de um grupo dedicado exclusivamente à preservação digital. A responsabilidade por essas atividades parece estar fragmentada entre as áreas de TI e operações, sem um foco institucional claro. Além disso, a alocação de orçamento para preservação digital não está formalizada e não existe um plano de desenvolvimento contínuo para os funcionários em relação a essas atividades.

B. Política e Estratégia: Políticas, estratégias e procedimentos que regulamentam a operação e o gerenciamento do arquivo digital[35].	1 – Consciência	O hospital não possui uma política ou estratégia formal implementada. A instituição pode ter algumas políticas relacionadas, como a de gestão de dados ou segurança da informação, mas não há um direcionamento específico para a preservação digital. Nesse contexto, as políticas existentes não abordam integralmente os desafios relacionados à preservação de longo prazo de ativos digitais. Além disso, a estratégia e os procedimentos para garantir o gerenciamento adequado e o acesso contínuo ao conteúdo digital não estão formalizados nem documentados.
C. Base Legal e Ética: Gerenciamento de direitos e deveres legais, sociais e culturais, conformidade com regulações relevantes e aderência à códigos de ética relacionados à aquisição, preservação e fornecimento de acesso aos objetos digital.	2 – Básico	No hospital é realizado um gerenciamento básico de direitos e deveres legais, sociais, culturais e éticos relacionados ao conteúdo digital. No contexto das regulamentações de saúde, o hospital segue as principais diretrizes da LGPD, bem como os códigos de ética do CFM, que regem a confidencialidade, integridade e acesso aos prontuários eletrônicos e outros dados clínicos dos pacientes. Apesar de ainda haver lacunas na implementação completa de políticas, o hospital já identificou e documentou as partes envolvidas com direitos legais e éticos, como pacientes e seus representantes legais, e adere ao Código de Ética Médica para assegurar a privacidade dos dados sensíveis de saúde. Além disso, já existem modelos de acordos legais e licenças necessários, como termos de consentimento dos

[35] Um arquivo digital nesse contexto se refere a um local de armazenamento para a guarda a longo prazo de dados digitais que não são necessários imediatamente, mas que ainda são considerados importantes.

		pacientes. O hospital também segue condutas profissionais básicas e está em conformidade com as regulamentações de preservação de dados médicos exigidas pela Lei nº 13.787/2018. No entanto, ainda carece de uma revisão e gestão contínua dos riscos legais e éticos mais abrangentes, que poderia ser mais bem documentado e monitorado de forma contínua.
D. Capacidade de TI: Competências da Tecnologia da Informação para apoiar as atividades de preservação digital.	0 – Consciência mínima	Atualmente, na Unidade de Saúde São Lucas, existe uma consciência mínima da necessidade de uma infraestrutura adequada de Tecnologia da Informação (TI) para apoiar as atividades de preservação digital. As práticas de suporte técnico, como monitoramento contínuo, atualizações de *software* ou implementação de ferramentas específicas de preservação digital, são mínimas ou inexistentes, o que limita a capacidade de gerenciar os objetos digitais em longo prazo.
E. Melhoria Contínua: Procedimentos para a avaliação da capacidade atual de preservação digital, a definição de metas e o monitoramento de progresso.	2 – Básico	Recentemente, a instituição realizou uma avaliação inicial de sua coleção de dados e capacidades de preservação digital, o que resultou na identificação de lacunas do contexto institucional e possibilidade de melhorias. Embora tenha sido identificado a necessidade de avançar em questões especificas, ainda não foram estabelecidos objetivos concretos nem um processo formal para guiar o desenvolvimento e avanço da maturidade de preservação digital.
F. Comunidade: Engajamento e contribuição com a vasta comunidade de preservação digital.	1 – Consciente	Existe um entendimento interno de que a troca de conhecimentos e o engajamento com a comunidade de preservação digital são passos importantes para melhorar suas práticas de preservação a longo prazo, ao compartilhar boas práticas e experiências. No entanto, o hospital ainda não se envolveu diretamente em

		redes ou eventos relacionados à preservação digital.
Capacidades de Serviço		
	Nível de Maturidade	**Justificativa**
G. Aquisição, transferência e ingestão: Processos para aquisição ou transferência de conteúdo e ingeri-los em um arquivo digital.	1 – Consciente	A Unidade de Saúde São Lucas tem consciência da necessidade de adquirir e transferir conteúdo digital para um arquivo digital, no entanto ainda não desenvolveu ou implementou um processo formal e documentado para essas atividades. Além disso, o hospital carece de uma área dedicada, física ou virtual, para realizar atividades de ingestão, como verificação de integridade de arquivos, identificação de formatos ou checagem de vírus.
H. Preservação de fluxo de *bits*: Processos para garantir o armazenamento e integridade do conteúdo digital a ser preservado.	2 – Básico	Algumas práticas já foram aplicadas, são elas: um regime de *backup*, c*hecksums* são gerados para garantir a integridade e a instituição tem uma noção de quais membros da equipe são autorizados a acessar determinados conteúdo.
I. Preservação de conteúdo: Processos para preservar o significado, usabilidade e funcionalidade do conteúdo digital ao longo do tempo.	2 – Básico	Os formatos dos arquivos foram identificados e já uma compreensão de quem são os usuários dos conteúdos digitais e como esses dados podem ser utilizados no futuro, principalmente em função das obrigações legais e de acessos frequentes a prontuários e dados clínicos.
J. Gerenciamento de metadados: Processos para criar e manter metadados suficientes para fornecer suporte à preservação, descobrimento e uso do conteúdo digital preservado.	2 – Básico	O hospital utiliza modelos para organizar e descrever as informações clínicas, como informações sobre o paciente (nome, idade, antecedentes clínicos), datas de atendimento, diagnóstico, tratamento e histórico médico. Além disso utiliza o modelo Subjetivo, Objetivo, Avaliação, Plano (SOAP) para manter o controle

		da evolução do paciente e incluem informações sobre data de consultas e exames, lista de diagnósticos e condições e detalhes sobre o plano de cuidados. Outros padrões terminológicos também são utilizados, como o CID-10, SIGTAP e CIAP-2. No que diz respeito à padrões de metadados, apenas o padrão DICOM para exames de imagens é adotado pela instituição.
K. Descoberta e Acesso: Processos para permitir a descoberta do conteúdo digital e fornecer acesso aos usuários.	2 – Básico	O acesso aos sistemas e dados do hospital é de acordo com as permissões e funções dentro da instituição. O hospital segue as principais diretrizes da LGPD e códigos de ética do CFM.

Fonte: elaborado pela autora baseado no DPC RAM (2024)

A avaliação revela um cenário ainda em estágio inicial de conscientização das capacidades de preservação digital. Conforme observado na Tabela 5.6, as capacidades de serviço demonstram uma maior consistência, ainda que em nível básico, e sugere uma base operacional mais bem estabelecida. No entanto, é importante ressaltar que as práticas identificadas não têm como foco principal a preservação digital, mas sim abordagens gerais de gestão de dados.

Além disso, as capacidades organizacionais apresentam maior variação, com destaque para a Capacidade de TI (nível 0), que representa um gargalo significativo para o desenvolvimento das ações de preservação digital na instituição. Áreas como Viabilidade Organizacional, Comunidade e Base Legal e Ética (todas nível 1) também requerem atenção para fortalecer a estrutura institucional. Este cenário indica a necessidade de um plano de ação focado principalmente no desenvolvimento das capacidades organizacionais, com ênfase especial no fortalecimento da infraestrutura de TI e na consolidação das estratégias institucionais.

5.1.3 Escopo e objetivos

Definir o escopo e os objetivos do plano de preservação digital é necessário para orientar as ações e garantir o sucesso das iniciativas na Unidade de Saúde São Lucas.

O escopo delimita as fronteiras do plano, especificando quais tipos de dados e sistemas serão contemplados, enquanto os objetivos estabelecem as metas a serem alcançadas, alinhadas com as necessidades institucionais e regulatórias.

O principal propósito desse plano é preservar os objetos digitais da Unidade de Saúde São Lucas. Devido ao rápido crescimento e a importância desses objetos, é imprescindível que se estabeleça uma estrutura para assegurar a autenticidade, disponibilidade e confiabilidade contínua dos dados clínicos, administrativos e operacionais ao longo do tempo. Isso inclui a preservação de prontuários eletrônicos, imagens médicas, resultados de exames, registros financeiros e outros dados essenciais para as operações da instituição. A preservação adequada desses dados garante a continuidade do atendimento, a conformidade com as regulamentações vigentes e a proteção das informações sensíveis dos pacientes.

O escopo do plano de preservação digital abrange todos os sistemas de informação utilizados pela Unidade de Saúde São Lucas, assim como todos os grupos de objetos digitais. Os objetivos principais do plano são:

1. Colaborar com diversas áreas da instituição, incluindo setores clínicos e administrativos, para alcançar as metas de preservação digital;
2. Adotar padrões, melhores práticas e requisitos regulatórios;
3. Comprometer-se com o treinamento sobre estratégias e ferramentas de preservação digital, desenvolvendo a expertise da equipe e a capacidade institucional em preservação de longo prazo;
4. Estabelecer processos e procedimentos para apoiar as atividades de preservação digital, maximizando o uso dos recursos disponíveis e garantindo a sustentabilidade das ações de preservação no futuro;
5. Identificar os itens físicos e digitais que devem ser preservados, aplicando as estratégias apropriadas para diferentes tipos de conteúdo, incluindo prontuários eletrônicos, imagens médicas, vídeos de telemedicina e registros administrativos.
6. Garantir o acesso seguro e contínuo aos materiais digitais, respeitando os direitos de propriedade intelectual e as normas de privacidade e confidencialidade dos dados de saúde;

7. Revisar e avaliar regularmente as políticas e procedimentos de preservação digital, levando em conta as mudanças na tecnologia, nos recursos disponíveis e nas necessidades institucionais.

A seleção é necessária porque o grande aumento no volume de documentos digitais eleva o custo de preservação, sendo que não são todos os materiais que precisam ser mantidos por tempo indeterminado. Assim, é importante a definição de critérios de seleção, requisitos de autenticidade e as prioridades da instituição com relação ao seu acervo digital. Como resultado, esse processo deve apresentar os objetos digitais selecionados, autênticos e em um formato de acordo com o previsto. (MOREIRA, 2017). Para isso, as seguintes perguntas devem ser respondidas:

1. Existe alguma política institucional que exija a preservação?
2. Há exigências legais ou normativas que obriguem a preservação?
3. O objeto pode ser facilmente substituído ou recuperado?
4. O custo para preservar é justificado em relação ao valor do objeto?

Se todas as respostas forem "não", a preservação pode não ser necessária. No entanto, se qualquer uma das respostas for "sim", devem ser avaliados fatores limitantes, como:

5. A instituição dispõe dos recursos necessários para preservação (espaço de armazenamento, capacidade técnica, *hardware*)?
6. Podem ser obtidos metadados suficientes do objeto?
7. O objeto contém informações sensíveis ou confidenciais?
8. Existe orçamento suficiente para preservar o objeto a longo prazo?

A resposta afirmativa a qualquer uma dessas perguntas indica que é necessário avaliar os fatores correspondentes como: viabilidade técnica; metadados de preservação; direitos autorais e recursos financeiros, para então definir o plano de preservação.

5.2 Gestão de Riscos

A gestão de riscos é uma etapa que permite avaliar as ameaças e vulnerabilidades associadas à manutenção de objetos digitais a longo prazo. Esta etapa envolve o processo de identificação, análise e mitigação de riscos que podem comprometer a integridade, acessibilidade e segurança dos dados clínicos, administrativos e financeiros gerados pela instituição.

5.2.1 Avaliação dos riscos

Avaliar os riscos significa realizar uma análise das ameaças potenciais que os objetos digitais podem enfrentar durante seu ciclo de vida. Essa análise deve considerar tanto fatores técnicos quanto organizacionais e ambientais e que possam impactar a preservação dos dados. No contexto da Unidade de Saúde São Lucas, são utilizados dispositivos de armazenamento como discos rígidos, unidades de estado sólido (SSDs), fitas magnéticas, discos ópticos e unidades removíveis, que podem se degradar com o tempo e resultar na perda ou corrupção de dados.

As causas dessa deterioração incluem tanto fatores físicos, como desgaste mecânico e exposição a condições ambientais inadequadas, quanto problemas lógicos, como corrupção de dados. A vida útil dos dispositivos também pode variar significativamente. Por exemplo, discos rígidos podem durar entre 5 a 8 anos, enquanto SSDs podem ter uma vida útil de 10 anos ou mais, mas seu desempenho pode ser afetado pela quantidade de dados gravados ao longo do tempo (PADRÃO, 2020). Temperaturas elevadas podem reduzir drasticamente a vida útil dos discos rígidos; portanto, é essencial controlar o ambiente e garantir que as ventoinhas estejam funcionando corretamente e não estejam obstruídas por poeira.

Fitas magnéticas, usadas para *backup*, têm maior durabilidade — 30 anos ou mais — se armazenadas em condições ideais de temperatura e umidade controladas; porém, são vulneráveis à desmagnetização e falhas nos drives de leitura (ELIAS, 2015). Novas tecnologias de fita magnética oferecem armazenamento de altíssima densidade. Tecnologias de cartuchos de alta densidade, como SDLT, LTO e AIT, são consideradas as mais confiáveis. No entanto, novas gerações de formatos de fita

aparecem regularmente, e a compatibilidade retroativa geralmente é oferecida apenas para 1 ou 2 gerações anteriores (KENNEY et al., 2003).

Discos ópticos, como CDs e DVDs, ainda utilizados principalmente para distribuição de *softwares* específicos da unidade de saúde, são ainda mais suscetíveis a danos físicos, como arranhões e oxidação, embora possam ter uma vida útil de até 50 anos se armazenados de forma adequada (SPASOJEVIC, 2024). Já as unidades removíveis, como *pendrives*, são mais práticas, mas apresentam risco de perda física e falhas inesperadas por desgaste ou corrupção de dados.

Além disso, esses dispositivos podem se tornar obsoletos com o tempo. Por exemplo, leitores de CDs e DVDs estão cada vez menos presentes em computadores novos, e a falta de suporte a essas tecnologias em novos dispositivos coloca em risco o acesso a dados armazenados nesses formatos. A obsolescência tecnológica não afeta apenas o *hardware*; os formatos de arquivos também podem se tornar obsoletos.

Um formato de arquivo é uma forma padrão de codificar informações para armazenamento em um arquivo de computador e especifica como os *bits* são usados para codificar informações em um meio de armazenamento digital (ARTEFACTUAL SYSTEMS INC, 2024). À medida que os s*oftwares* passam por atualizações frequentes, a compatibilidade com arquivos criados em versões mais antigas costuma ser descartada. Isso significa que arquivos que não foram migrados para novas versões correm o risco de não serem lidos pelo *software* atualizado. Além disso, as versões antigas do *software* podem não estar mais disponíveis ou podem não funcionar em computadores modernos ou em versões atuais do sistema operacional (KENNEY et al., 2003).

Embora seja incomum para formatos amplamente utilizados, formatos menos populares podem deixar de ser suportados e se tornar ilegíveis com o passar dos anos. Isso ocorre também quando o *software* necessário para ler esses arquivos não é mais mantido ou atualizado, forçando a migração para novos formatos compatíveis ou o risco de perda de acesso (DPC, 2015). Para considerar os riscos associados com os formatos, é realizado, inicialmente, o processo de Identificação, que analisa as informações fornecidas sobre um arquivo para então definir seu formato (ARTEFACTUAL SYSTEMS INC, 2024). No contexto da Unidade de Saúde São Lucas, os formatos identificados incluem:

Tabela 5.7 – Formatos utilizados

Formato	Uso
.dcm/.dicom	Imagens médicas
.csv	Relatórios gerados pelos sistemas
.pdf	Relatórios gerados pelos sistemas e documentos administrativos
.doc/.docx/.odt	Documentos administrativos
.mp4, .avi, .wav	Arquivos de telemedicina e gravações de consulta
.xlsx	Planilha de dados
.txt	Textos simples
.svs/.tiff /.ndpi	Imagens digitalizadas de lâminas digitais
.png/.jpg	Imagens não categorizadas

Fonte: elaborado pela autora

O formato DICOM (*.dcm*), utilizado para imagens médicas, pode se tornar um problema se versões mais antigas perderem compatibilidade com atualizações mais recentes do *software* utilizado para armazenar esses arquivos. Além disso, o PEC, sistema utilizado para prontuários eletrônicos dos pacientes, gera relatórios em formatos *csv* e *.pdf*. Documentos administrativos são salvos em *.doc*/*.docx* e *.pdf*, e podem gerar dificuldades de acessibilidade a longo prazo se armazenados com recursos avançados ou criptografados. Para os arquivos multimídia de vídeo e áudio são utilizados os formatos *.mp4*, *.avi* e *.wav* em arquivos de telemedicina e gravações de consultas, podem se tornar ilegíveis caso *codecs* ou *players* deixem de ser suportados.

O suporte ao formatos *.svs* e *.ndpi* também pode ser interrompido se os visualizadores de imagens específicos se tornarem obsoletos ou incompatíveis com versões futuras de *software*. Adicionalmente aos arquivos *doc*/*.docx* e planilhas no formato *.xlsx*, dependem de softwares proprietários para serem lidos corretamente, o que cria uma dependência que eleva o risco de obsolescência e inutilização dos dados. Esse risco ocorre porque os detentores da licença controlam o uso da tecnologia, podendo restringir o acesso de terceiros no presente ou futuro. Esses formatos proprietários geralmente só podem ser acessados pelo *software* original ou por

aplicativos licenciados, não compartilham especificações técnicas nem propriedade intelectual e, por não terem requisitos de licenciamento abertos ao público, dificultam a preservação a longo prazo dos dados. (QUEENSLAND GOVERNMENT, 2021).

As especificações dos formatos definem a subdivisão adequada, a codificação, a sequência, o arranjo, o tamanho e as relações internas que identificam exclusivamente um formato específico e permitem que ele seja interpretado e exibido corretamente. A especificação do formato *.tiff*, por exemplo, define o bloco básico de construção de um arquivo e seu comprimento máximo, e em seguida detalha, *byte* por *byte*, a estrutura interna válida. Um arquivo que não cumpra exatamente esses requisitos pode ser irreconhecível ou renderizado incorretamente se for lido por um programa que interpreta *Tagged Image File Format* (TIFF) (KENNEY et al., 2003).

Além desses fatores, a integridade dos dados dos pacientes é fundamental para a correta continuidade de tratamentos e para a conformidade com exigências legais. Diversos riscos específicos podem comprometer essa integridade, como falhas nos sistemas, que podem resultar em perda ou corrupção dos dados devido a problemas técnicos, ou incompatibilidades de *software*. Erros humanos na inserção ou atualização de informações, ataques cibernéticos, e a integração inadequada dos diferentes sistemas utilizados também representam ameaças. A instituição ainda depende de fornecedores externos para serviços de TI, o que pode introduzir vulnerabilidades adicionais.

É imprescindível que seja feita a interpretação correta das informações dos pacientes, e isso pode depender de informações adicionais que estavam implícitas no contexto original de criação ou recebimento, mas que se tornam menos claras ao serem revisitadas posteriormente. Nesse sentido, é destacado a importância da criação de metadados junto à criação do dado, uma vez que os metadados auxiliam na gestão e compreensão da informação.

A criação dos dados na Unidade de Saúde São Lucas ocorre tanto através do sistema Prontuário Eletrônico do Cidadão (PEC) quanto por meio de fichas de Coleta de Dados Simplificada, utilizadas como alternativa manual em situações específicas. Essas fichas são empregadas para registrar cadastros, visitas domiciliares, atendimentos e outras atividades realizadas em contextos em que há falta de conectividade ou interrupção no fornecimento de energia elétrica. Posteriormente, os dados coletados manualmente são digitados no sistema eletrônico. No entanto, o

processamento das fichas é feito de forma automática e exige que o servidor do sistema fique ligado, conectado à internet, durante a noite.

As informações são inseridas no PEC seguindo o modelo de Registro Clínico Orientado por Problemas (RCOP), composto por quatro componentes: Base de Dados, Lista de Problemas (Folha de Rosto), Evolução – que utiliza o método Subjetivo, Objetivo, Avaliação e Plano (SOAP) – e Folha de Acompanhamento (Fichas de resumo e fluxograma). Base de Dados inclui informações essenciais como identificação, antecedentes pessoais e familiares, além de exames clínicos e fatores de risco, com atualizações periódicas para mapear a situação geral do paciente. A Lista de Problemas é dividida entre problemas ativos, que requerem tratamento contínuo, e problemas latentes, que não necessitam de tratamento imediato, mas precisam ser acompanhados, como o histórico de tabagismo ou doenças familiares.

O componente Evolução SOAP organiza as notas de atendimento de forma estruturada em quatro partes: Subjetivo (relato do paciente), Objetivo (observações clínicas), Avaliação (inferências e diagnóstico) e Plano (condutas e tratamentos recomendados). Finalmente, a Folha de Acompanhamento é utilizada para monitorar a evolução de doenças crônicas e tratamentos, incluindo a frequência de sintomas e resultados de exames complementares.

Esses procedimentos clínicos são mapeados utilizando SIGTAP, CIAP-2 e, em alguns casos, CID-10. Também é realizada a identificação de agravos de notificação compulsória, como doenças infecciosas, que podem ser geradas diretamente pelo PEC. No entanto, o sistema apenas gera relatórios em formato PDF, que são enviados manualmente às autoridades responsáveis, já que o PEC não realiza a integração automática com a base nacional do Sistema de Informação de Agravos de Notificação (SINAN).

Para exames de imagem, a instituição utiliza o *Picture Archive and Communication System* (PACS). No entanto, não existe interoperabilidade entre esses sistemas (PEC e PACS). A Unidade de Saúde São Lucas utiliza o sistema SISREG para realizar a gestão hospitalar e funcionalidades administrativas. Entre os padrões de interoperabilidade dos sistemas, podemos citar o HL7 e o openEHR.

O sistema SISREG permite ainda gerenciar o fluxo de agendamentos de consultas e exames especializado, possibilitando tanto que a instituição receba e organize as

demandas encaminhadas por outras unidades de saúde da rede pública, quanto encaminhe os dados do paciente caso identifique a necessidade de um atendimento ou procedimento especializado após a consulta inicial. Os atendimentos são priorizando conforme a gravidade e a disponibilidade de vagas.

Com o aumento do volume de exames médicos, tele consultas e prontuários eletrônicos, a demanda por capacidade de armazenamento também cresce rapidamente. No entanto, instituições de saúde pública, como a Unidade de Saúde São Lucas, frequentemente operam com orçamentos restritos, o que impacta diretamente na capacidade de implementar práticas de preservação. Custos com armazenamento, infraestrutura de TI, renovação de sistemas obsoletos, e a treinamento de pessoal são alguns dos fatores que geram maior pressão financeira.

Diante dessa realidade, torna-se inviável armazenar todos os dados gerados de forma indefinida, o que leva à necessidade de seleção e descarte de informações. No entanto, na Unidade de Saúde São Lucas, não há um processo formal de seleção ou descarte, e todos os dados gerados acabam sendo armazenados. GRÁCIO (2012) destaca que a decisão de preservação e descarte deve ser pautada por políticas claras de avaliação documental, que levem em consideração o valor legal, administrativo, fiscal e histórico dos documentos. A preservação de dados médicos envolve, além de questões técnicas, aspectos éticos e legais, como a garantia da privacidade e do sigilo dos pacientes, e a correta manutenção de registros de saúde que possam ser utilizados em casos futuros.

5.2.2 Estratégia de preservação

No cenário atual da Unidade de Saúde São Lucas, algumas ações e estratégias devem ser consideradas para implementação, a fim de mitigar os riscos identificados associados à perda ou inconsistência das informações. É importante que as estratégias de preservação digital considerem, para cada tipo de objeto digital, o método mais adequado a ser utilizado, seja na preservação do conteúdo ou na sua forma física. O foco é garantir tanto a autenticidade quanto a integridade do objeto, de modo a assegurar sua acessibilidade ao longo do tempo (GRÁCIO, 2012).

Embora a expectativa de vida útil das mídias de armazenamento seja uma referência útil para prever sua durabilidade, a realidade pode diferir significativamente. Uma das principais causas de falhas prematuras nas mídias de armazenamento é o armazenamento inadequado. Manter temperaturas moderadas, próximas de 20°C, e umidade relativa em torno de 40%, bem como evitar grandes e rápidas variações desses parâmetros, são práticas recomendadas para prolongar a vida útil das mídias. O ambiente deve estar livre de poeira, com pressão positiva para evitar a entrada de partículas, e evitar a exposição a campos magnéticos, vapores e luz ultravioleta. O manuseio deve ser feito de forma correta: uso de luvas sem fiapos, evitar tocar diretamente na superfície de CDs, guardar a mídia adequadamente e limitar o acesso apenas a pessoal treinado são algumas das recomendações para garantir a conservação adequada dos objetos digitais.

Ainda que procedimentos de *backups* sejam incluídos na rotina da instituição, eles não atendem aos requisitos mais específicos para garantir a preservação a longo prazo dos materiais digitais. O armazenamento para preservação exige um nível mais alto de redundância geográfica, recuperação de desastres mais forte, planejamento de longo prazo e o monitoramento ativo da integridade dos dados. Isso permite a detecção antecipada de problemas (DPC, 2015).

Entre as boas práticas destacadas nos Níveis de Preservação da NDSA, é recomendado a implementação de:

a) Múltiplas cópias independentes dos objetos digitais;

b) Cópias geograficamente separadas em locais diferentes;

c) Diferentes tecnologias de armazenamento nas cópias;

d) Combinação de técnicas de armazenamento *online* e *offline*;

e) Monitoramento ativo do armazenamento para garantir que quaisquer problemas sejam detectados e corrigidos rapidamente.

Outros fatores relacionados à deterioração das mídias de armazenamento e a obsolescência de *hardware* e *software* reforçam que "[...] as mídias são suportes transitórios que prestam sua função somente por um período limitado de tempo e que a transferência para novas mídias é absolutamente necessária" (THOMAZ; SOARES, 2004 *apud* GRÁCIO, 2012). Na Unidade de Saúde São Lucas, boa parte

dos servidores e dispositivos de armazenamento em uso já ultrapassaram sua vida útil recomendada. Dessa forma, é necessária a implementação de um processo periódico de refrescamento e migração. Ambas as estratégias exigem o monitoramento constante dos profissionais designados para que as atualizações necessárias sejam realizadas antes que os dados fiquem comprometidos.

A migração "é a estratégia de preservação mais aplicada, tanto em contextos institucionais, como no domínio doméstico" (LEE *et al.*, 2002). No entanto, depende de diversos critérios para que seja implementada com sucesso em uma instituição, que inclui a experiência técnica da equipe envolvida, as expectativas dos usuários, o orçamento disponível, os equipamentos adequados e o tempo disponível para a execução das atividades (MOREIRA, 2017). A migração de *hardware* envolve a transição de dados e sistemas de um ambiente de *hardware* antigo para um novo, garantindo que todos os componentes funcionem corretamente no novo sistema. No entanto, na prática, é reconhecido que as restrições orçamentárias dificultariam a aquisição de novos equipamentos e a atualização da infraestrutura existente. Além disso, o processo de migração demanda tempo, planejamento detalhado e recursos humanos qualificados, o que pode sobrecarregar as equipes já limitadas do hospital.

Embora o refrescamento de suporte seja uma alternativa menos custosa em comparação à migração de *hardware*, ele também enfrenta desafios semelhantes. A limitação orçamentária continua sendo um obstáculo, pois o refrescamento requer a substituição periódica das mídias de armazenamento e a alocação de recursos para gerenciar e monitorar o processo. Além disso, mesmo transferindo os dados para mídias mais atuais, o refrescamento não resolve problemas como a obsolescência dos equipamentos e sistemas de *hardware*, que ainda precisam ser atualizados.

Apesar dessas dificuldades, as estratégias se alinham com as prioridades estabelecidas no Plano de Ação da ESD28, que enfatiza a modernização da infraestrutura tecnológica e a melhoria da qualidade dos serviços de saúde. Esse alinhamento reforça a importância estratégica dessa iniciativa e possibilita que a instituição busque parcerias e apoio institucional para viabilizar a atualização necessária.

No que diz respeito à migração de *softwares* ou atualização de versão, a estratégia inclui garantir que os dados e processos sejam compatíveis com outros *softwares* ou com versões mais recentes dele, evitando riscos de incompatibilidade. No caso do sistema de prontuário eletrônico utilizado na Unidade de Saúde São Lucas, ele é

mantido por uma equipe de desenvolvimento parceira, e a cada nova versão, o Ministério da Saúde fornece orientações sobre os passos necessários para a migração. Esses passos incluem a verificação do ambiente, avaliação da versão atual e do banco de dados, além da realização de *backup* dos dados, entre outras medidas para garantir uma transição segura.

Os mesmos procedimentos de atualização e migração devem ser aplicados aos sistemas PACS e AGHU. Considerando que as atualizações para esses sistemas são lançadas também por diferentes grupos de desenvolvimento, deve-se estabelecer um plano estruturado para realizar a atualização de versões para garantir o funcionamento ideal e a compatibilidade entre eles. Esse plano deve incluir a equipe responsável pela execução das atualizações, um cronograma detalhado, os procedimentos de *backup* e recuperação de dados, além das estratégias de comunicação com os usuários para minimizar impactos nas operações da instituição. Também é importante considerar os profissionais que estarão envolvidos e a documentação de todo o processo.

Embora as estratégias de emulação e encapsulamento sejam reconhecidas no campo da preservação digital e estejam entre as mais implementadas nas organizações, sua aplicação na Unidade de Saúde São Lucas não seria a mais adequada. A emulação exige um alto nível de conhecimento técnico e recursos financeiros para desenvolver e manter emuladores ao longo do tempo. Isso pode ser inviável para a instituição, considerando as restrições orçamentárias e a não garantia de total integridade e funcionalidade dos sistemas. Já o encapsulamento pode não ser prático devido ao grande volume, à diversidade de formatos de dados utilizados pela instituição e por depender da implementação de outras estratégias, como a emulação ou pelo desenvolvimento de conversores e visualizadores (SANTOS; FLORES, 2015). Além disso, também não resolve problemas de obsolescência de *hardware* e pode dificultar a interoperabilidade com sistemas futuros. Os mesmo problemas são encontrados na preservação/conservação da tecnologia.

Outro ponto destacado na análise de riscos se refere à obsolescência de formatos. Na prática, o problema pode não ser tão grave quanto a comunidade de preservação digital previa há cerca de 20 anos. Muitos formatos de arquivo estabelecidos ainda estão em uso, ainda são suportados e continuam funcionais (DPC, 2015). No entanto, com a evolução das tecnologias, pode haver uma redução gradual no

suporte a formatos mais antigos, tornando os dados armazenados nesses formatos suscetíveis à inacessibilidade. Para diminuir esse risco, é recomendado o uso de padrões abertos (SANTOS; FLORES, 2017; ARELLANO, 2008; GRÁCIO, 2012).

Formatos de código aberto são bastante populares por serem livres de restrições proprietárias, o que dá aos usuários uma sensação de maior controle e autonomia ao utilizá-los (DPC, 2015). O JPEG 2000, por exemplo, é um padrão ISO (ISO/IEC 15444) que tem sido amplamente utilizado em várias áreas, incluindo a imagiologia médica. Quando integrado ao padrão DICOM, permite a compressão das imagens sem comprometer de forma significativa a qualidade. Considerando a necessidade de compartilhamento e acesso a longo prazo de imagens médicas na Unidade de Saúde São Lucas, a adoção do padrão JPEG 2000 integrado ao DICOM, já utilizando na instituição, seria altamente recomendada.

Outras iniciativas para definição de formatos abertos incluem o padrão *Portable Document Format – Archival* (PDF/A) e o *Open Document Format* (ODF), ambos regulamentados pela ABNT NBR ISO 19005-1 e ISO 26300, respectivamente. O formato PDF/A tem o objetivo de preservar a aparência visual de documentos eletrônicos ao longo do tempo, independente dos sistemas utilizados para sua criação, armazenamento ou leitura. Além disso, busca facilitar o acesso futuro e atender a possíveis necessidades de migração, por meio da incorporação de metadados e da definição da estrutura lógica e das propriedades semânticas dos documentos. Assim, é considerado mais adequado à preservação a longo prazo do que o PDF tradicional (SULLIVAN, 2006).

Já o padrão ODF é um formato de arquivo baseado em XML usado para armazenamento e troca de documentos de texto, planilhas de cálculo, apresentações, bancos de dados e desenhos vetoriais. A adoção desses formatos se mostra uma boa estratégia para a preservação digital na Unidade de Saúde São Lucas.

O uso de padrões abertos é mais adequado considerando a questão de direitos autorais, normalmente podem ser usados e modificados por qualquer pessoa sem restrições, e facilita a compreensão da estrutura dos formatos de arquivo e funcionamento dos *softwares*, uma vez que esse tipo de formato e é desenvolvido em um processo público e colaborativo, com especificações e propriedade intelectual disponíveis abertamente. Além disso, simplifica o processo de preservação, especialmente no que se refere ao uso de plataformas de *hardware* e *software*, já que

são suportados por uma ampla variedade de softwares ou são independentes de plataforma e permite migração entre diferentes ambientes técnicos sem ficar restrito a um único fornecedor, e ainda contribui para a redução de gastos com licenças de *softwares* proprietários. (SANTOS; FLORES, 2018; QUEENSLAND GOVERNMENT, 2021).

Dentre as opções de formatos disponíveis, devem ser priorizados aqueles que oferecem maior expectativa de acesso a longo prazo, com formatos amplamente utilizados sendo mais confiáveis. Um resumo dos formatos mais adequados foram definidos na Tabela 5.8.

Tabela 5.8 – Resumo dos formatos dos arquivos para preservação

Tipo	**Formato**	**Especificações**
Imagem médica	DICOM	ISO 12052:2017
Relatórios	CSV	RFC 4180
Documentos e Relatórios administrativos	PDF/A PDF/A-1 PDF/A-2 PDF/A-3	ISO 19005-1:2005 ISO 19005-2:2011 ISO 19005-3:2012
Arquivos de telemedicina e Gravações	Broadcast WAV MP4 (sem compressão)	EBU Tech 3285 ISO/IEC 14496-14
Planilhas de dados	Open Document Spreadsheet (ODS)	ISO/IEC 26300
Textos simples	TXT (sem formatação)	Nenhuma base padronizada
Imagens Digitalizadas de Lâminas	TIFF	Adobe Systems Incorporated
Imagens Não Categorizadas	TIFF JPEG com metadados EXIF	ISO/IEC 10918-7:2023 ISO/IEC 14495

Fonte: adaptado de SIEBRA *et al* (2018)

Essa etapa é a Normalização, definida como o processo de pegar um arquivo de um determinado formato e transformá-lo em outro formato para uma finalidade específica, como acesso ou preservação - por exemplo, o Archivematica pode conter

regras para converter um arquivo PNG em JPG para acesso e um TIFF para preservação (ARTEFACTUAL SYSTEMS INC, 2024). Para apoiar esse processo, outras ferramentas especializadas também podem ser integradas ao fluxo de trabalho. O ImageMagick convert[36] permite a conversão entre vários formatos gráficos, incluindo o formato DICOM, que pode ser convertido para formatos mais comuns para fins de acesso ou visualização. O ffmpeg[37] é utilizado em arquivos audiovisuais; e o Ghostscript[38] e o ps2pdf[39] são usados para transformar materiais em PDF.

No entanto, alguns formatos proprietários específicos, como os gerados por *softwares* utilizados na farmácia e laboratório de análises clínicas, não podem ser facilmente migrados para outros formatos devido a questões de compatibilidade e às funcionalidades especializadas que esses sistemas oferecem. Ainda que especificações proprietárias e fechadas representem alguns dos *softwares* mais duradouros em uso, elas também tendem a evoluir mais rapidamente e a existir em muitas versões diferentes e com compatibilidade limitada a versões anteriores. Portanto, esses formatos são os mais vulneráveis à obsolescência, pois enfrentam o risco duplo de mudanças rápidas na especificação e de estarem atrelados a um único produto ou empresa (KENNEY et al., 2003).

Tanto as técnicas de refrescamento quanto a de conversão (migração de formato) envolvem procedimentos que podem expor os objetos digitais a riscos de alteração, por isso, deve-se fazer o uso de ferramentas para garantir a integridade e a autenticidade desses objetos ao longo do processo (SANTOS; FLORES, 2015b). DELANEY; JONG (2015) afirmam que, para a preservação digital, esses são dois conceitos chave: a integridade, que significa que o conteúdo permanece inalterado não foi corrompido ao longo do tempo de preservação, e a autenticidade, que significa que o conteúdo é o que afirma ser.

---

[36] Disponível em https://imagemagick.org/script/convert.php

[37] Disponível em https://www.ffmpeg.org/

[38] Disponível em https://www.ghostscript.com/

[39] Disponível em https://www.ps2pdf.com/

Por isso, também deve ser incluído o processo de Validação. Essa etapa verifica se os arquivos não estão corrompidos, que seguem corretamente as especificações do formato e que contêm todos os elementos necessários (ARTEFACTUAL SYSTEMS INC, 2024). Ferramentas como o JHOVE (para PDF, JPEG e TIFF), veraPDF (especificamente para PDF/A), MediaConch (para arquivos audiovisuais) e jpylyzer (para imagens no formato JPEG 2000) automatizam a validação de arquivos digitais.

Os autores também destacam ainda que a integridade e a autenticidade são garantidas pelas estratégias, ações e fluxos de trabalho da preservação pelos quais o conteúdo passa, assim como pelo registro sistemático de metadados ao longo de seu ciclo de vida. SALES (2022) propõe um modelo (Figura 5.1) específico de metadados para preservação de PEP, fundamentado em padrões amplamente reconhecidos como: DC; *Metadata Object Description Schema* (MODS); PREMIS; ANSI/NISO Z39.87; *Encoded Archival Description* (EAD); e *National Library of New Zealand* (NLNZ) *Schema*.

Figura 5.1 – Modelo de Metadados para Preservação Digital dos PEPs

1. PRONTUÁRIO
1.1 Identificador
1.1.1 Número do prontuário Master
1.1.2 Número do prontuário AGHUX
1.2 Data do documento
1.2.1 Criação
1.2.2 Última modificação
1.3 Nome da instituição
1.3.1 Caráter
1.3.2 Localização
1.4 Identificação do paciente
1.4.1 Nome
1.4.1.1 Nome social
1.4.2 Data de nascimento
1.4.3 CPF
1.4.4 Nome da mãe
1.4.5 Credo
1.4.6 Histórico de doenças
1.4.6.1 CID-10
1.4.7 Data
1.4.7.1 Internação
1.4.7.2 Alta
1.5 Formulários
1.5.1 Anamnese
1.5.1.1 Categoria
1.5.1.2 Data
1.5.1.3 Encaminhamento
1.5.1.4 Diagnóstico
1.5.1.5 Doença
1.5.1.5.1 CID-10
1.5.2 Evolução Clínica
1.5.2.1 Tipo
1.5.2.2 Categoria profissional
1.5.2.3 Data
1.5.2.4 Diagnóstico
1.5.2.5 Doença
1.5.2.5.1 CID-10
1.5.2.6 Medicamento
1.5.2.6.1 Nível de risco
1.5.2.7 Tratamento
1.5.3 Registro de Controle
1.5.3.1 Sinais vitais
1.5.3.2 Balanço hídrico
1.5.3.3 Protocolos
1.5.4 Prescrições
1.5.4.1 Prescrição médica
1.5.4.2 Prescrição de enfermagem
1.5.5 Sumário de Alta
1.5.5.1 História clínica
1.5.5.2 Exame físico
1.5.5.3 Evolução e complicação
1.5.5.4 Diagnóstico definitivo
1.5.5.4.1 CID-10
1.5.5.5 Orientação médica
1.5.5.6 Condição de alta
1.6 Tamanho do arquivo
1.7 Certificação
1.8 Versão do Software
1.9 Custódia
1.9.1 Local
1.9.2 Custodiadores
1.9.3 História custodial
1.10 Condição de acesso
1.11 Proveniência
1.12 Item relacionado

2. EQUIPE MULTIPROFISSIONAL
2.1 Nome do profissional
2.2 Profissão
2.2.1 Especialidade
2.3 Identificador

3. EXAMES
3.1 Tipo de exame
3.2 Exames laboratoriais
3.2.1 Categoria
3.2.2 Formato
3.2.2.1 Conjunto de caracteres
3.3 Exames de imagem
3.3.1 Categoria
3.3.2 Resolução
3.3.3 Dimensões
3.3.4 Cores de referência
3.3.5 Bits
3.3.6 Tamanho
3.3.7 Data
3.3.8 Formato do arquivo
3.3.9 Metadados descritivos
3.3.10 Formato de armazenamento
3.3.11 Autenticação
3.3.12 Metadados de avaliação da imagem

4. MODIFICAÇÃO DE METADADOS
4.1 Registro de modificação
4.2 Item modificado
4.3 Data da modificação
4.4 Histórico de mudança

Fonte: SALES (2022)

O modelo é composto por quatro entidades que descrevem diferentes aspectos dos prontuários de pacientes. A Entidade 1. Prontuário inclui 11 elementos, como identificadores, datas e nomes (paciente e organização), além de documentos importantes para o histórico de saúde do paciente, como anamnese, evolução clínica e prescrições. Para preservar a integridade, autenticidade e interpretação do prontuário, é necessário registrar também dados sobre a certificação e versão do *software* e a custódia do prontuário ao longo do tempo (SALES, 2022). Os identificadores representados na Figura 5.1 são dos prontuários Master e AGHUX, prontuário analógico e eletrônico, respectivamente, utilizados na instituição de referência para a elaboração do modelo.

Conforme explica SALES (2022), a Entidade 2. Equipe Multiprofissional contém 3 elementos que registram a participação de diversos profissionais de saúde, como médicos, enfermeiros e técnicos, no prontuário, indicando nome, profissão e identificador, o que reforça a proveniência e a responsabilidade no atendimento ao paciente. A Entidade 3. Exames agrupa 3 elementos para armazenar resultados de exames laboratoriais e de imagem, adaptando-se a diferentes formatos, como imagem, áudio, vídeo e texto, e seguindo padrões técnicos como o DICOM. A Entidade 4. Modificação de Metadados registra o histórico de alterações nos metadados de preservação para manter uma trilha de auditoria, a fim de monitorar a integridade e assegurar o controle sobre cada modificação realizada.

No entanto, a incorporação desse modelo requer diversas adaptações. Podemos destacar principalmente a adoção de padronizações nas estruturas de representação de dados, mensagens e vocabulário, que se trata de um processo mais trabalhoso e lento devido à complexidade do setor; e a necessidade de garantir a segurança e confidencialidade dos dados dos pacientes durante a integração dos sistemas (SALES, 2022).

Esses metadados oferecem aos usuários uma maneira de gerenciar os objetos digitais e podem ser usados para auditoria, por exemplo, no sentido de rastrear o histórico do objeto e fornecer provas da origem da fonte, o que é importante para a vida útil do objeto. Dessa forma, os metadados são importantes para qualquer processo de preservação digital, "pois expressam todas as estratégias de preservação digital aplicadas ao objeto digital desde sua criação, bem como as informações necessárias para sua representação" (GRÁCIO, 2012).

Além disso, qualquer alteração nos objetos digitais deve ser documentada, detectável e gerenciável. A soma de verificação (ou *checksum*), uma espécie de "impressão digital" de um arquivo, são ideais para verificar se mudanças indesejadas ocorreram em objetos digitais. No entanto, às vezes esses objetos serão alterados de forma deliberada, por exemplo, se um formato de arquivo for migrado. Isso faz com que o *checksum* mude, exigindo que novas somas de verificação sejam estabelecidas após a migração, tornando-se a maneira de verificar a integridade dos dados do novo arquivo no futuro. Por isso, como afirma CAPLAN (2009), "uma soma de verificação armazenada como metadado pode ser usada para comprovar se um arquivo sofreu alguma alteração em um determinado intervalo de tempo".

No caso do PEC, há um módulo específico que faz auditorias nos registros realizados. Esta funcionalidade permite saber quais ações foram executadas dentro do sistema e quem as executou, a partir da geração de trilhas de auditoria. Adicionalmente, é recomendada a utilização periódica dessa funcionalidade para que seja possível corrigir erros de manuseio do prontuário e detectar eventuais irregularidades.

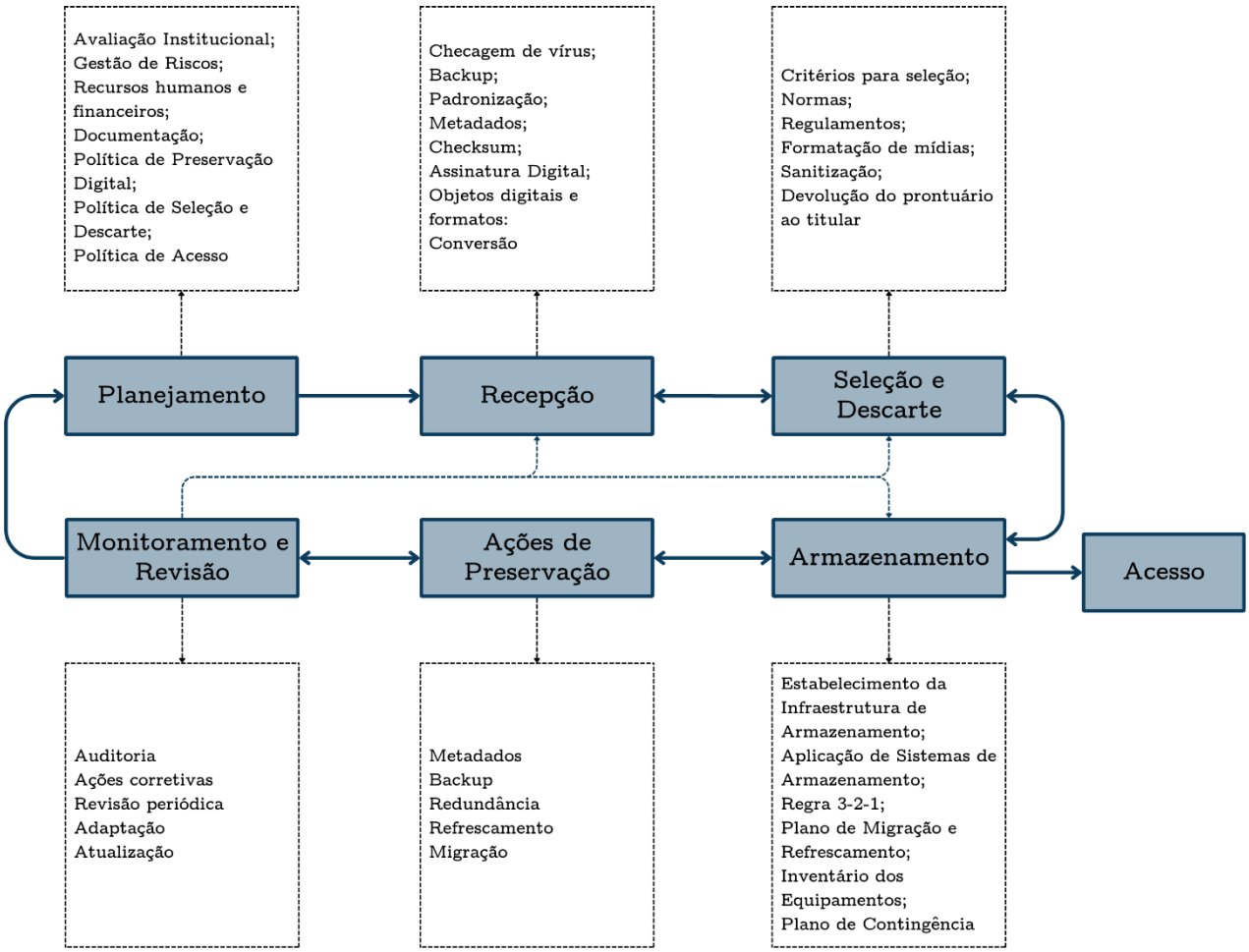
Com o volume de informações digitais produzidas pela instituição, é necessário selecionar os dados relevantes que devem ser preservadas. Portanto, surge um aspecto importante na preservação digital: a seleção (GRÁCIO, 2012). Os critérios definidos na Avaliação institucional auxiliarão na definição e na justificativa do que deve ser preservado e por quanto tempo preservar. Como explica MELLO (2020), é a partir desses critérios e de políticas, normas, regulamentos e atos administrativos, que a equipe encarregada da seleção avalia se a preservação desse objeto digital é de responsabilidade da instituição.

Assim como a seleção, a gestão do descarte é igualmente importante, e tem como objetivo determinar a duração da preservação de cada objeto digital e a forma como ele será descartado (GRÁCIO, 2012). No caso de instituições de saúde, os objetos digitais relacionados ao paciente devem ser mantidos de forma permanente, conforme o CFM. Caso seja necessário o descarte, a lei nº 13.787, define que o processo de eliminação deverá preservar a intimidade do paciente e o sigilo e a confidencialidade das informações, assim como a destinação final dos prontuários e sua eliminação registradas em forma de regulamento. Uma alternativa à eliminação completa, é a devolução do prontuário ao paciente. (BRASIL, 2018).

5.3 Plano de ação da Preservação digital

O plano de ação, ilustrado na Figura 5.2, elaborado para a Unidade de Saúde São Lucas foi estruturado considerando os componentes definidos pelo modelo OAIS para orientar as etapas de preservação digital da instituição. Além disso, baseou-se nas etapas e ações de preservação propostas pelo modelo do DCC.

Figura 5.2 – Plano de ação



Fonte: elaborado pela autora baseada nos modelos OAIS e ciclo de vida dos dados do DCC

As etapas do plano de ação são ilustradas por retângulos azuis, representando as fases das ações implementadas. Os fluxos entre as etapas são indicados por setas, mostrando a sequência de atividades, enquanto os retângulos pontilhados contêm as ações a serem realizadas em cada etapa, assim como os objetos a serem produzidos por elas. O plano foi modelado em sete etapas principais: *planejamento*, *recepção*, *seleção e descarte*, *armazenamento*, *acesso*, *ações de preservação*, e *monitoramento*

e revisão. A recomendação para começar com ações mais simples e adaptá-lo conforme a instituição consolida suas atividades foi mencionada por especialistas na área, como uma estratégia para garantir a sustentabilidade e consistência dos processos de preservação digital.

Ao adotar um sistema específico para a preservação digital, como o Archivematica, boa parte dessas atividades teriam maior nível de automação e controle, facilitando o cumprimento das etapas necessárias e reduzindo a necessidade de intervenções manuais. Dependendo da capacidade organizacional da instituição, o sistema permitiria configurar fluxos de trabalho automatizados para cada etapa, desde o planejamento até o monitoramento e revisão, garantindo que as ações fossem realizadas de maneira consistente e conforme os padrões definidos pela instituição.

5.3.1 Planejamento

A primeira etapa do modelo, o *planejamento*, está diretamente ligada aos objetivos e à cultura organizacionais, uma vez que os objetos digitais a serem preservados e disponibilizados dependem das informações produzidas pela instituição e de seus valores (GRÁCIO, 2012). As atividades dessa etapa devem ser realizadas de forma contínua durante todo o ciclo de vida do objeto digital, conforme orientado pelo DCC. Os modelos de curadoria digital do DataOne, da DigitalNZ e do UK *Data Archive* incluem essa etapa nos seus respectivos ciclos e a evidenciam como uma atividade transversal e contínua.

No plano de ação da Unidade de Saúde São Lucas, há um fluxo de retorno para a etapa de *planejamento*, o que enfatiza essa adaptação constante às mudanças tecnológicas, institucionais e de necessidades dos usuários, ainda que não deva ocorrer, obrigatoriamente, de forma sequencial. Para contextualizar o cenário desta pesquisa e fundamentar a elaboração do plano de ação, foram conduzidas as ações de avaliação institucional e gestão de riscos dessa etapa. A avaliação institucional buscou compreender as principais necessidades e objetivos específicos da instituição de saúde em relação à preservação digital, enquanto a gestão de riscos focou em identificar e analisar potenciais ameaças nesse cenário. Esse levantamento inicial permitiu delinear as etapas e ações preventivas e corretivas para o plano de ação.

Parte das ações já previamente executadas podem ser executadas por ferramentas que automatizam esses processos. No mapeamento dos recursos digitais da instituição, por exemplo, ferramentas como o DROID[40], desenvolvida pelo *The Nacional Archives* (UK), podem gerar metadados descritivos (nome do arquivo, tipo do arquivo, formato, versão, extensão), técnicos (tamanho, última modificação, caminho do arquivo no sistema e status da análise) e de preservação (somas de verificação, identificador PRONOM[41]) de maneira bem intuitiva e com ótima documentação de apoio aos usuários (DPC, 2015).

Além disso, outras ferramentas como o Siegfried[42] e o *Format Identification for Digital Object*s (FIDO)[43], ambas utilizadas pelo Archivematica, automatizam o processo de identificação de formatos de arquivos. O Siegfried é um identificador de formatos rápido que também utiliza o PRONOM para reconhecer uma ampla variedade de formatos digitais. Já o FIDO é uma ferramenta de linha de comando que também se baseia nas assinaturas do PRONOM para realizar a identificação de formatos. Essa automatização não apenas acelera os processos, mas também garante maior precisão na geração automática de metadados técnicos e de preservação.

Em complemento à essas ações, o planejamento também inclui definir as responsabilidades das equipes e dos indivíduos envolvidos e estimar os recursos financeiros necessários para implementar as políticas e a infraestrutura de gestão (GRÁCIO, 2012). Em um ambiente OAIS, os papeis são representados pelo *produtor*, *consumidor*, *gestor* (externos ao ambiente) e *administração do sistema* (interno ao ambiente), como descrito no tópico *2.6.1 – O modelo de referência Open Archival Information System (OAIS)*.

Na Unidade de Saúde São Lucas, tanto pacientes quanto profissionais de saúde desempenharão papeis múltiplos dentro desse ambiente de preservação. A

---

[40] Disponível em: https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/

[41] O PRONOM é um banco de dados mantido pelo The National Archives (UK) que fornece informações detalhadas sobre formatos de arquivos, incluindo identificadores e especificações.

[42] Disponível em https://www.itforarchivists.com/siegfried.

[43] Disponível em https://github.com/openpreserve/fido/releases.

participação dos usuários destaca ainda a importância de inseri-los na definição do que deve ser preservado (GRÁCIO, 2012). Além disso, outras instituições de saúde que compartilham ou consultam informações preservadas pela instituição também podem atuar como *produtores* e *consumidores*. Essas instituições podem, por exemplo, enviar registros para serem integrados a instituição ou consultar dados para continuidade de cuidados com pacientes.

Na *administração do sistema*, são realizadas, de fato, as ações de preservação (descritas na etapa seguinte), e esse componente requer uma definição clara das responsabilidades e tempo suficiente para sua realização. Faz parte da Política de Preservação Digital a atribuição de funções e responsabilidades e a identificação de como a organização proporcionará oportunidades de treinamento para que a equipe possa desenvolver, manter ou aprimorar sua *expertise* em preservação digital (BROWN, 2013). Para realizar ou apoiar o trabalho de preservação digital, são necessárias competências em diversas áreas específicas. MCMEEKIN; CURRIE (2022) resumem essas competências em cinco grupos. As responsabilidades e atividades específicas de cada área podem ser vistas na Tabela 5.9.

Tabela 5.9 – Competências para a preservação digital

Área de competência	**Responsabilidades**
Governança, Recursos e Gestão	Os profissionais precisam ser capazes de contextualizar a preservação digital dentro dos objetivos organizacionais, aplicar técnicas de gestão de riscos, e definir estratégias de planejamento que garantam a sustentabilidade e segurança das operações. As atividades incluem redigir políticas, desenvolver planos de continuidade e gerenciar contratos.
Comunicação e *Advocacy*	As atividades incluem identificar as necessidades dos usuários, promover a preservação digital na instituição e documentar procedimentos de forma clara e acessível. As atividades nesta área incluem a criação de documentos, realização de treinamentos e desenvolvimento de mensagens direcionadas para diferentes grupos de interesse.

Tecnologia da Informação	É fundamental que a equipe tenha conhecimentos técnicos, como conhecimento de conceitos de TI, programação, segurança da informação e desenvolvimento de fluxos de trabalho. Essas habilidades possibilitam a implementação e manutenção de sistemas de armazenamento e segurança necessários para a preservação digital. Atividades comuns incluem instalação de *software*, análise de segurança, e planejamento de infraestruturas de armazenamento.
Responsabilidades Legais e Sociais	Competências em conformidade regulatória, impacto social e ambiental, inclusão, diversidade e ética orientam a preservação digital dentro de um marco de responsabilidade social. Os profissionais devem estar preparados para aplicar princípios éticos e assegurar que as práticas de preservação digital respeitem as regulamentações legais e promovam a inclusão e diversidade.
Competências específicas do Domínio de Preservação Digital	Para garantir que os objetos digitais sejam preservados e acessíveis, são necessárias habilidades em gestão de metadados, princípios de gestão da informação, planejamento e implementação de ações de preservação, estratégia e gestão organizacional e acessibilidade e suporte ao usuário.

Fonte: adaptado de MCMEEKIN; CURRIE (2022)

No entanto, é importante considerar que, em instituições públicas como a Unidade de Saúde São Lucas, onde os recursos podem ser limitados, pode não ser viável estabelecer cargos dedicados para a preservação digital. Essa realidade exige uma abordagem flexível, em que os profissionais conciliam suas responsabilidades de preservação digital com outras atividades institucionais. Isso, de acordo com BROWN (2013), não representa um problema, desde que as responsabilidades sejam claramente definidas e que haja tempo suficiente para a execução das atividades de

preservação. Além disso, a instituição deve garantir que suas atividades de preservação digital sejam realizadas por pessoal suficiente e com as habilidades adequadas.

Na Política de Preservação Digital uma estimativa dos custos também deve ser elaborada. De acordo com Grácio (2012), questões econômicas representam um dos principais desafios da preservação digital, principalmente devido aos custos envolvidos. Por isso, toda atividade relacionada à preservação digital deve ser previamente planejada e avaliada para que possa ser executada posteriormente. Por isso, as partes envolvidas devem se comprometer a garantir os recursos necessários para a continuidade das atividades a longo prazo (SANTOS; FLORES, 2015a).

De acordo com FORMENTON; GRACIOSO (2020), além dos recursos humanos e da implantação, operação e manutenção das atividades de preservação, fatores como os recursos materiais e a missão e os objetivos institucionais, incluindo o tipo e volume das coleções, os níveis de preservação e acesso definidos, e o período proposto para as ações, também impactam diretamente nos custos a serem analisados. Por exemplo, o custo de manutenção do ambiente de preservação do Arquivo Nacional é estimado em R$ 2.300.000,00 por ano (dois milhões e trezentos mil reais), sem incluir os custos de aquisição de equipamentos e de implantação dos serviços (FARIA; SILVA, 2024).

Alguns modelos genéricos de preservação digital, como o *LIFE3 Costing Model* (LIFE3), o *Total Cost of Preservation* (CDL-TCP), o *NASA Cost Estimating Tool* (NASA-CET), o *Cost Model for Digital Preservation* (CMDP), entre outros, fornecem boas ferramentas para a estimativa de custo nas diferentes etapas da preservação digital. BOTE; FEIJOO; RUIZ (2013) propõe um modelo para o cálculo de custos para a preservação digital em organizações de saúde ou terceiros que mantêm RES, visto na Figura 5.3. No entanto, os métodos citados são carentes de estudos na Ciência da Informação nacional (FORMENTON; GRACIOSO, 2020).

No modelo de custo de Bote, Feijoo e Ruiz (2013), cada "seção" é uma fase do projeto de preservação digital separada para facilitar o controle de custos. Para calcular os custos indiretos – aqueles que não estão ligados diretamente a uma atividade específica –, o processo segue quatro etapas:

1. Alocação Primária: Primeiro, são identificados todos os custos de cada seção. Cada custo recebe um critério de alocação, que indica como ele será distribuído entre as seções que o utilizaram. É importante também classificar esses custos como fixos ou variáveis;
2. Reclassificação dos Custos Fixos: Com essa classificação pronta, é possível seguir para a próxima fase, que envolve recalcular os custos fixos de forma mais detalhada para cada seção. Esse cálculo dos custos fixos indiretos considera o "nível normal de atividade" de cada seção, ou seja, o quanto uma seção consegue produzir em condições normais, com base na sua capacidade;
3. Alocação Secundária: Nesta fase, os custos indiretos são redistribuídos entre as diferentes seções, de acordo com o critério definido anteriormente;
4. Atribuição Final do Custo: Por fim, os custos são atribuídos a cada seção com base em uma unidade de medida (como horas de trabalho, volume de uso, etc.), adaptando os critérios de alocação de acordo com a realidade de cada organização (BOTE; FEIJOO; RUIZ, 2013).

Como explicam os autores, esse método distribui os custos fixos e variáveis de forma proporcional, o que é importante em setores de tecnologia onde os custos fixos são significativos. Esse modelo segue uma rotina repetitiva e predefinida e nem sempre existe um fim definitivo do processo (os RES precisam ser mantidos indefinidamente, por exemplo). A fórmula abaixo representa o modelo de custo total do processo descrito:

$$\text{Custo} = \text{CEN} + \sum_{i=1}^{n} \text{CPRE} + \text{CAC} + \text{CEX}$$

Onde:

- CEN (Custo de Entrada) é o custo de registrar um novo conjunto de RES no sistema. É calculado com base nos custos diretos de trabalho (como o tempo gasto pelos funcionários para inserir os dados) e nos custos indiretos das seções envolvidas no processo de entrada.
- CPRE (Custo de Preservação) é o custo de manter os dados armazenados de forma segura ao longo do tempo. É calculado periodicamente (por exemplo,

mensalmente), somando o consumo de recursos em cada seção que participa da preservação.

- CAC (Custo de Recuperação e Acesso) é o custo que se refere ao processo de buscar e acessar registros específicos quando solicitado.
- CEX (Custo de Saída) é o custo de remover ou arquivar informações que chegaram ao fim de sua vida útil, com base na data de expiração.

Para diminuição nesses custos, GRÁCIO (2012) sugere a participação em programas de cooperação, formação de parcerias e a união de experiências com outras instituições, o que leva ainda a outros fatores como a diminuição dos esforços de preservação, troca de conhecimentos e possibilita uma melhor interoperabilidade dos objetos digitais.

Figura 5.3 – Modelo de Custo para Preservação Digital de RES

Seções Conceito	Critério de Alocação	Variável / Fixo	Recepção	Digitalização	Pré-Ingestão	Ingestão	Armazenamento	Auditoria	Migração	Recuperação e Acesso	Destruição	Total
Alocação primária total			TR(V+F)	TDig(V+F)	TP-I(V+F)	TI(V+F)	TSC(V+F)	TA(V+F)	TM(V+F)	TRA(V+F)	TDes(V+F)	$\sum$
Custo variável primário total			TRV	TDigV	TP-IV	TIV	TSCV	TAV	TMV	TRAV	TDesV	
Custo fixo primário total			TRF	TDigF	TP-IF	TIF	TSCF	TAF	TMF	TRAF	TDesF	
Atividade normal de cada seção			Número normal de novas solicitações	NHPI Normal	NHC Normal	NHPI Normal	Tamanho Normal em KB	NHA Normal	NHPI Normal	Número normal de solicitações * Tamanho	Número normal de destruição * Tamanho	
Custo unitário fixo (I)			TRF/Número normal de novas solicitações	TDigF/NHPI Normal	TP-IF/NHC Norma	TIF/NHPI Normal	TSCF/Tamanho Normal em KB	TAF/NHA Normal	TMF/NHPI Normal	TRAF/Número normal de solicitações * Tamanho	TDesF/Número normal de destruição. * Tamanho	
Atividade real no período (II)			Número de novas solicitações	NHPI	NHC	NHPI	Tamanho em KB	NHA	NHPI	Número de solicitações * Tamanho	Número de destruição * Tamanho	
Recalculo do custo fixo primário total (I)*(II)			TRFr	TDigFr	TP-IFr	TIF	TSCFr	TAFr	TMFr	TRAF	TDesFr	
Custo primário total após o recálculo do custo fixo (III)			TRV + TRFr = TRFV	TDigV + TDigFr = TDigFV	TP-IV + TP-IFr = TP-IFV	TIV + TIF = TIFV	TSCV + TSCFr = TSCFV	TAV + TAFr = TAFFV	TMV + TMFr = TMFV	TRAV + TRAFr = TRAFV	TDesV + TDesFr = TDesFV	
Atribuição do custo												
Unidades de atividade (IV)			Número de novas solicitações	NHPI	NHC	NHPI	Tamanho em KB	NHA	NHPI	Número de solicitações * Tamanho	Número de destruição * Tamanho	
Custo por unidade de atividade (III)/(IV)			TRFV/Novas solicitações	TDigFV/NHPI	TP-IFV/NHC	TIFV/NHPI	TSCFV/KB	TAFFV/NHA	TMFV/NHPI	TRAFV/Número de solicitações * Tamanho	TDesFV/Número de destruição * Tamanho	

Fonte: traduzido de BOTE; FEIJOO; RUIZ (2013)

No planejamento também devem ser elaboradas as políticas de seleção, descarte e acesso. Os critérios de seleção foram pré-definidos na Avaliação Institucional, no entanto, a instituição deve levar em consideração alguns riscos associados à essa etapa. Objetos infectados por *malware* precisam ser desinfetados antes da recepção; mídias danificadas podem exigir ferramentas especializadas para recuperação dos dados; e mídias obsoletas ou formatos incomuns podem necessitar de *hardware*, *software* ou serviços específicos para serem acessados. A documentação insuficiente é outro risco, pois a falta de informações pode afetar a autenticidade dos objetos digitais. Além disso, grandes volumes de dados podem exceder a capacidade de armazenamento disponível o que implica em custos adicionais e desafios técnicos (BROWN, 2013).

Já a Política de Descarte precisa determinar as condições e os processos para o descarte dos objetos digitais, respeitando os princípios de retenção e descarte adequado conforme a legislação. A Lei nº 13.709 determina que os dados pessoais podem ser eliminados a qualquer momento por requisição do titular ou em caso de infração às normas da própria LGPD. Além disso, a instituição deve informar imediatamente a terceiros com quem compartilhou os dados sobre qualquer eliminação realizada, para que eles repliquem o mesmo procedimento.

Os processos definidos nessa etapa devem garantir que o descarte seja realizado de forma segura e irreversível, evitando qualquer possibilidade de recuperação não autorizada dos dados eliminados. Por fim, a Política de Acesso, conforme orientado por Hedstrom (1997), deve balancear o acesso à informação e a proteção dos direitos dos usuários e dos titulares dos dados. Isso deve incluir a definição de diferentes níveis de acesso e a aplicação de medidas de segurança para proteger a integridade e a confidencialidade dos dados.

5.3.2 Recepção

Essa etapa envolve tanto a criação de novos objetos digitais quanto o recebimento de terceiros. É de interesse da instituição que o conteúdo seja completo, preciso e utilizável, tanto para os objetivos atuais quanto para as necessidades futuras. Para isso, essa etapa considera a definição de *softwares* e formatos que sejam amplamente

suportados (conforme definido na Gestão de Riscos), adoção de convenções consistentes para nomes de arquivos, implementação de práticas de armazenamento e *backup* adequadas e planejadas transições futuras. No Archivemativa ou em qualquer ambiente OAIS, essas ações se configuram no PSI, e devem começar já no momento da criação e recepção dos materiais, com orientações claras sobre como nomear arquivos, armazenar dados de forma segura e seguir as obrigações legais e práticas recomendadas (DPC, 2015).

Como recomenda o DPC em relação à nomeação de arquivos, é preferível utilizar nomes curtos e descritivos que incluam o conteúdo e a data, fornecendo um contexto claro e compreensível tanto para humanos quanto para sistemas computacionais. Deve-se evitar o uso de espaços ou caracteres especiais (exceto "-" ou "_"), pois isso minimiza o risco de interpretações incorretas por *hardware* ou *software*. A data deve seguir o padrão ISO 8601:2004 no formato "YYYY-MM-DD", o que permite uma rastreabilidade consistente de versões. Para indicar diferentes versões de um arquivo, deve ser adotado uma convenção, por exemplo, um identificador de versão, como "v1", "v2" ou "v_final", conforme necessário.

A nomeação com significado facilita a localização do arquivo e ajuda a identificar seu conteúdo. No entanto, essa identificação completa é mais complexa e outras informações que descrevem as propriedades ou atributos devem ser atribuídas aos materiais digitais (INTERPARES PROJECT, 2011). O DCC (2019) destaca que, ao produzir objetos digitais, podem ser atribuídos metadados arquivísticos, administrativos, descritivos, estruturais ou técnicos. Esse é o processo de Caracterização. Recomenda-se que os metadados de preservação sejam inseridos já no momento da criação (MELLO, 2020). No PEC (sistema de prontuário eletrônico) e no PACS (sistema de imagiologia médica), a adoção de um padrão de metadados de preservação, como o modelo proposto por SALES (2022), permite padronizar a descrição e o gerenciamento dos objetos digitais.

Nessa etapa, também é essencial garantir a autenticidade do objeto, utilizando assinatura digital ou soma de verificação para assegurar sua integridade. A Lei nº 14.063 estabelece a obrigatoriedade de assinatura digital com o certificado digital Infraestrutura de Chaves Públicas Brasileira (ICP-Brasil) (também chamada de assinatura eletrônica qualificada) do profissional de saúde para laudos médicos, atestados e prescrições. Para garantir a validade jurídica e a autenticidade dos

prontuários e outros documentos médicos ao serem armazenados, é necessário utilizar o tipo adequado de assinatura eletrônica conforme previsto na lei (BRASIL, 2020).

A verificação da integridade de arquivos também pode ser realizada de forma simples e automática com o uso de somas de verificação (ou *checksums*). De forma mais específica, um *checksum* é um valor gerado com base no conteúdo de um arquivo, utilizando um algoritmo matemático. Esses algoritmos são desenvolvidos para que qualquer alteração mínima, como a mudança de um único *bit* no arquivo, produza um valor totalmente diferente. Ao calcular a soma de verificação de cada arquivo em um PSI antes de sua transferência ao armazenamento, é possível testar a integridade do arquivo posteriormente: recalculando a soma e comparando com o valor original, assim qualquer divergência apontará alterações no arquivo.

A criação de *checksums* durante o processo de recepção é útil para estabelecer uma referência de integridade para o futuro. Entre os algoritmos de *checksum* mais utilizados estão o MD5 e o SHA-1 (BROWN, 2013); entre as principais ferramentas, estão o *Checksum By Corz*[44] (ferramenta gratuita) e o *Fixity Pro*[45] (ferramenta desenvolvida especialmente para a comunidade de preservação digital) (DPC, 2015).

Como destaca MELLO (2020), "outro ponto importante relaciona-se ao controle de possíveis vírus que poderão danificar as informações registradas". Para evitar que essas ameaças causem danos aos arquivos e ao repositório como um todo, é essencial implementar um procedimento de quarentena como primeiro passo ao receber um novo objeto. Esse processo deve incluir tanto a detecção e contenção de ameaças quanto o tratamento ou descarte de conteúdo infectado, em uma área de quarentena isolada do repositório principal. (BROWN, 2013)

---

[44] Disponível em: https://corz.org/windows/software/checksum/

[45] Disponível em: https://www.fixitypro.com/

5.3.3 Seleção e Descarte

A Unidade de Saúde São Lucas mantem a maioria de seus objetos digitais de forma permanente, com exceção de documentos administrativos e financeiros e imagens de lâminas utilizadas nos laboratórios da instituição. Essa seleção é realizada oficialmente, pautada pela legislação que regulamentam os dados de instituições brasileiras de saúde. Todavia, conforme apontado por MELLO (2020), pode ser necessário estabelecer uma escala de prioridades para a preservação, especialmente considerando o grande volume de informações digitais em uma instituição de saúde, o que pode acarretar lentidão nos processos ou impor restrições orçamentárias e técnicas, dificultando parcial ou totalmente a execução completa do processo.

No caso dos dados enviados ao DATASUS, eles permanecem armazenados nas bases de dados do Ministério da Saúde pelo período necessário para apoiar a formulação e execução de políticas públicas de saúde. Esses dados pessoais passam por processos de anonimização. Para os objetos que não estão diretamente vinculados à informações dos pacientes, serão aplicados, de maneira sistemática, os critérios de seleção previamente estabelecidos neste plano.

Para validar as informações contidas nos prontuários do paciente, a Resolução CFM nº 1821/2007 tornou obrigatória a criação da Comissão de Revisão de Prontuários nas instituições de saúde, que é responsável por assegurar que: (I) os prontuários contenham todos os elementos obrigatórios; (II) estejam em conformidade com os padrões estabelecidos, tanto em formato eletrônico quanto em papel; (III) sejam devidamente preenchidos, armazenados e manuseados (CFM, 2007).

Além disso, conforme discutido nos capítulos anteriores, o período de preservação de um objeto digital deve atender às exigências legais estabelecidas por leis, normas e portarias, além de considerar a frequência de uso, no caso de objetos que não possuem restrições legais específicas. Para o descarte, a definição dos objetos deve incluir a criação de uma tabela de temporalidade e a metodologia para execução dessa atividade. Essa tabela precisa ser estabelecida com base nos objetivos da instituição e nessas exigências legais, além de considerar soluções tecnológicas que previnam o acúmulo de lixo digital (GRÁCIO, 2012).

Grácio (2012) também destaca que a instituição pode adotar dois procedimentos para gerenciar objetos digitais selecionados para descarte: a eliminação definitiva dos arquivos ou a transferência para uma estrutura designada para objetos descartados, que não seriam incluídos nas estratégias de preservação, mas podem retornar à infraestrutura de preservação e acesso, caso necessário. Esse processo de descarte deve iniciar ou pela detecção automática através de metadados ou pela verificação da Política de Descarte. Para a eliminação de dados, as estratégias devem ser realizadas de maneira segura para evitar a recuperação de informações sensíveis e estar em conformidade com as diretrizes da LGPD.

As estratégias de eliminação, tecnicamente conhecidas como sanitização de dados, precisam ser adaptadas de acordo com o meio de armazenamento. Por exemplo, para os registros em papel, a trituração transversal é mais indicada do que a trituração em tiras, pois isso dificulta a reconstrução das folhas destruídas. Em registros eletrônicos, a simples exclusão ou formatação das mídias não se mostra eficaz, pois ainda permite a recuperação completa dos dados.

Para eliminar permanentemente os dados, técnicas como a desmagnetização, onde mídias de armazenamento magnético, como discos rígidos, fita magnética ou disquetes, são expostas a um campo eletromagnético com intensidade superior à sua blindagem, o que permite alterar a estrutura natural da mídia a ponto de inutilizá-la; e a sobregravação, onde sequências repetitivas de zeros e uns são gravadas em todo o disco, sobrescrevendo os dados originais, são recomendadas. Nessa situação, *softwares* dedicados para esse propósito devem ser utilizados (PCPD, 2011).

A destruição física também continua sendo uma estratégia segura para eliminar objetos digitais. Métodos como desintegração, incineração, esmagamento, trituração e derretimento podem ser empregados de acordo com a necessidade (PCPD, 2011). Quando a destruição é solicitada pelo titular dos dados, todas as cópias dos dados pessoais devem ser eliminadas, cabendo à instituição informar as entidades públicas e privadas com as quais compartilhou os dados.

No entanto, a LGPD prevê exceções para a completa eliminação dos dados, em que é autorizada a retenção mesmo diante da solicitação pelo titular, são elas: (I) cumprimento de obrigação legal ou regulatória; (II) estudo por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais; (III) transferência a terceiro, desde que respeitados os requisitos de tratamento de dados dispostos na

lei; e (IV) uso exclusivo pela própria instituição, com proibição de acesso por terceiros, e desde que anonimizados os dados (BRASIL, 2018a).

O processo de anonimização significa remover dos dados pessoais qualquer informação que possa identificar ou associar os dados, de forma direta ou indireta, a um indivíduo (BRASIL, 2018a). A Tabela 5.10 resume algumas técnicas básicas de anonimização. Cada método possui uma aplicação específica, dependendo do tipo de dado e do grau de anonimização desejado.

Tabela 5.10 – Técnicas básicas de anonimização

Técnica de Anonimização	**Descrição**
Supressão de Registros	Consiste em remover completamente um registro que contenha dados facilmente identificáveis ou que possam ser considerados *outliers*. A exclusão desses registros é útil para prevenir a identificação de indivíduos únicos no conjunto de dados. Por exemplo, em um estudo sobre pacientes com determinada doença rara, registros de pacientes com características únicas podem ser removidos para evitar que sejam reconhecidos com base em suas particularidades.
Mascaramento de Caracteres	No mascaramento, alguns caracteres de uma informação são substituídos por símbolos, como "*" ou "x", para ocultá-los parcialmente. Esse método é útil em dados como números de telefone ou e-mails. Por exemplo, o número de telefone "555-123-4567" pode ser exibido como "555-**-****", protegendo parcialmente a identidade da pessoa.
Pseudonimização	A pseudonimização substitui dados identificáveis por valores fictícios, que

	podem ser únicos para cada indivíduo, tornando a reidentificação mais difícil. Ela pode ser irreversível (descartando os valores originais) ou reversível, caso a instituição precise reidentificar os dados. Por exemplo, em um banco de dados de licenças de condução, nomes reais podem ser substituídos por códigos alfanuméricos. "João Silva" pode se tornar "ID12345", com o mapeamento original guardado de forma segura para permitir reidentificação, se necessário.
Generalização	Esta técnica reduz a precisão de um dado, mantendo-o útil para análises sem revelar informações exatas. Por exemplo, em vez de registrar a idade exata “24 anos”, o dado pode ser generalizado para a faixa etária “21-30 anos”. Em uma pesquisa sobre pacientes, endereços completos podem ser substituídos por nomes de ruas sem números residenciais, de modo que apenas a localização aproximada seja compartilhada.
Perturbação de Dados	Modifica levemente os valores de dados numéricos, como altura ou peso, tornando-os menos precisos. Isso é feito por meio de arredondamentos ou introdução de ruído aleatório, sem comprometer a validade para análises agregadas. Em um estudo sobre a relação entre peso e altura de pacientes, valores de altura podem ser arredondados para o múltiplo de cinco mais próximo, transformando, por

	exemplo, “173 cm” em “175 cm”.
Troca de Dados	Nesta técnica, valores de certos atributos são trocados entre registros, mantendo a representatividade dos dados, mas removendo o vínculo direto com indivíduos específicos. Em um conjunto de dados de exames médicos, o diagnóstico de um paciente pode ser trocado com o de outro, de modo que as informações sejam mantidas para análise estatística, mas dificultando a reidentificação de um paciente específico.
Agregação de Dados	Quando o estudo não requer registros individuais, os dados podem ser convertidos para valores agregados, como médias ou totais, para análises populacionais. Em uma análise de doações para uma instituição de caridade, ao invés de manter o histórico de cada doador, os dados são agrupados em faixas de renda, indicando apenas o total doado por cada faixa.

Fonte: PDPC (2022)

Entre os *softwares* para anonimização de dados, o Eclipse[46] é destinado exclusivamente para dados de saúde, e permite que grandes volumes de informações sensíveis sejam transformados de maneira automatizada e segura. Com essa ferramenta, é possível anonimizar também imagens médicas (no padrão DICOM), relatórios de saúde e dados não estruturados. No entanto, por ser uma ferramenta paga, a implementação do Eclipse pode representar uma barreira para a instituição.

[46] Disponível em: https://privacy-analytics.com/eclipse-software/

Outro ponto importante dessa etapa é que haja um registro documentado de todas as atividades de descarte, detalhando datas, responsáveis e os meios utilizados, para fins de auditoria e conformidade legal. A instituição também deve assegurar que todos os funcionários e colaboradores envolvidos estejam cientes da Política de Descarte e recebam treinamento adequado para executá-la corretamente. Implementar controles internos e realizar revisões periódicas dos processos de descarte contribuirá para identificar possíveis falhas e aprimorar as práticas existentes.

No caso de devolução do prontuário ao paciente, deve ser entregue pessoalmente ou através de meio digital seguro. Considerando a seleção e o descarte em uma instituição como a Unidade de Saúde São Lucas, em que os objetivos e lideranças podem mudar ao longo do tempo, deve ser reavaliado continuamente o que deve ser preservado e o que pode ser descartado (GRÁCIO, 2012).

5.3.4 Armazenamento

Com a infraestrutura técnica e de pessoal já estabelecida nas etapas anteriores e ajustada às necessidades específicas da instituição e aos requisitos de preservação dos objetos digitais, é possível dar início às atividades de armazenamento desses objetos. O objetivo principal dessa etapa é inserir o objeto digital na infraestrutura tecnológica (abrangendo *hardware*, *software* e formato) mais adequada para atender às demandas da instituição e dos usuários em relação à busca, recuperação, acesso e preservação desse objeto (GRÁCIO, 2012).

No contexto do modelo de referência OAIS, essa etapa envolve o recebimento dos PAIs gerados na etapa de Recepção. Esses pacotes contêm a informação que deve ser preservada, acompanhada por um conjunto completo de metadados necessários e informações descritivas para apoiar os serviços de preservação e acesso dentro do sistema OAIS (LIRA; SIEBRA, 2021). Essa etapa envolve ainda definir especificações físicas e lógicas sobre como esses dados serão registrados e mantidos em um suporte adequado.

Os sistemas de armazenamento utilizados da Unidade de Saúde São Lucas foram projetados para os objetos digitais que estão em uso ativo. Como recomendado por GRÁCIO (2011), a estrutura de armazenamento deve, preferencialmente:

- Utilizar sistemas com alta capacidade de armazenamento e dispositivos de acesso de alta velocidade;
- Manter uma estrutura de *backup* confiável;
- Ter um sistema de redundância de banco de dados e de *hardware*;
- Ter um sistema de detecção e recuperação automática de falhas;
- Manter uma estrutura de redes de computadores adequada para acesso dos usuários aos sistemas;
- Dispor de sistemas de armazenamento com mecanismos de segurança;
- Ter acesso restrito aos equipamentos e aos objetos digitais.

No entanto, para definição dessa estrutura, é preciso considerar "a infraestrutura da instituição, o conhecimento técnico dos profissionais envolvidos, os recursos financeiros disponíveis e o contexto tecnológico da época, buscando a melhor relação custo-benefício" (GRÁCIO, 2012). Como destacado pelo Arquivo Público de São Paulo (2022), o uso de tecnologias e soluções de armazenamento precisa ser realizado de forma correta para que possam oferecer bons níveis de segurança de dados, acesso rápido ao conteúdo quando necessário e custos alinhados ao orçamento planejado. Para isso, é exigido um planejamento e gerenciamento cuidadosos, que considerem não apenas as demandas atuais da instituição, mas também a evolução tecnológica e as futuras necessidades de preservação.

Entre os princípios para o armazenamento para a preservação digital, uma prática amplamente recomendada é a implementação da regra 3-2-1. Essa estratégia consiste em manter três cópias dos dados: duas em dispositivos de armazenamento diferentes e uma terceira cópia fora do local. A regra não só ajuda a diminuir os riscos de perda de dados como também aumenta a redundância, o que é fundamental em situações de falhas ou desastres. O uso combinado de sistemas de armazenamento *online* e mídias *offline*, assim como a utilização de diferentes tipos de tecnologia de armazenamento, promove ainda a diversidade e segurança dos dados (DPC, 2015).

É recomendado ainda o uso de somas de verificação e outras medidas de fixidez para registrar e monitorar periodicamente a integridade de cada cópia digital. No caso de detecção de corrupção ou perda de dados, uma das cópias deve ser utilizada para criar uma substituição. Adicionalmente, é essencial armazenar as informações de fixidez tanto junto aos materiais digitais quanto em sistemas separados, o que adiciona uma camada extra de segurança. Outros aspectos incluem a observação contínua das tecnologias e fornecedores, a avaliação de riscos e migrações proativas.

O setor de armazenamento digital é notoriamente dinâmico, com tecnologias, produtos e serviços apresentando ciclos de vida relativamente curtos. Por isso, o monitoramento contínuo das tecnologias em uso permite que a instituição identifique o momento certo de realizar migrações, evitando que os materiais digitais fiquem em risco. Além disso, é necessário acompanhar a viabilidade dos fornecedores de armazenamento, garantindo que as soluções adotadas continuem seguras e viáveis ao longo do tempo (DPC, 2015).

Reduzir a variedade de tipos de mídia legada e minimizar o número de sistemas usados simplifica o processo de gestão dos dados e facilita o controle e a manutenção deles. A documentação adequada de como os materiais foram recebidos, transferidos e configurados nos sistemas de armazenamento é igualmente importante, pois contribui para a criação de trilhas de auditoria que garantem a autenticidade dos dados, oferecendo credibilidade e segurança às práticas de preservação.

A equipe deve elaborar a documentação e elaboração de um inventário dos equipamentos disponíveis, identificando informações como modelo, fabricante, data de aquisição, capacidade de armazenamento, conteúdo armazenado, permissões de acesso, prazo de garantia e previsão de descomissionamento. Esse inventário deve incluir diagramas das conexões de rede e dos circuitos elétricos envolvidos, seguindo padrões de notação e símbolos reconhecidos e documentados (ARQUIVO PÚBLICO DE SÃO PAULO, 2022).

Essa atenção detalhada à infraestrutura pretende prevenir as diversas ameaças que podem comprometer a preservação adequada dos objetos digitais. Essas incluem as ameaças físicas, como a instabilidade do material, armazenamento inadequado (incluindo temperatura, umidade, luz e poeira), desgaste por uso excessivo (principalmente para mídias que utilizam contato físico), desastres naturais (como incêndios, enchentes e terremotos), falhas de infraestrutura (problemas de

encanamento, sistemas elétricos ou de controle climático); ameaças tecnológicas, que são as causadas problemas no *hardware* ou *software*; e ameaças humanas, que são os erros humanos, falta de profissionais especializados, e até atos de sabotagem, como roubo e vandalismo.

Por isso, além de manter um ambiente de armazenamento adequado, o planejamento da infraestrutura deve contar ainda com um plano e infraestrutura de contingência, que assegure o fornecimento contínuo de energia elétrica, estabeleça protocolos de resposta a situações de sinistro (como falhas catastróficas de equipamentos) e contemple um plano de investimento contínuo para migração de *hardware* e refrescamento. (ARQUIVO PÚBLICO DE SÃO PAULO, 2022) Entretanto, o desenvolvimento desse plano não deve ser tratado como um processo único e definitivo; ele precisa ser constantemente testado e ajustado conforme mudanças nas circunstâncias, como a introdução de novos equipamentos, reorganização de instalações físicas ou a chegada de novos funcionários.

Para os objetos digitais já preservados e armazenados, será necessária, ao longo de seu ciclo de vida, a aplicação de estratégias de preservação digital que mantenha o objeto disponível. (GRÁCIO, 2012).

5.3.5 Ações de preservação

As ações de preservação aplicadas devem abranger aspectos relacionados ao *hardware*, *software*, *formatos* e suportes dos objetos digitais. Nesse contexto, a equipe deve:

- Considerar o cenário atual da instituição;
- Conhecer as novas tecnologias de informação e comunicação e suas tendências;
- Compreender os objetos digitais e as alterações registradas nos metadados de preservação;
- Levar em conta tanto a infraestrutura de preservação quanto o ambiente de acesso;

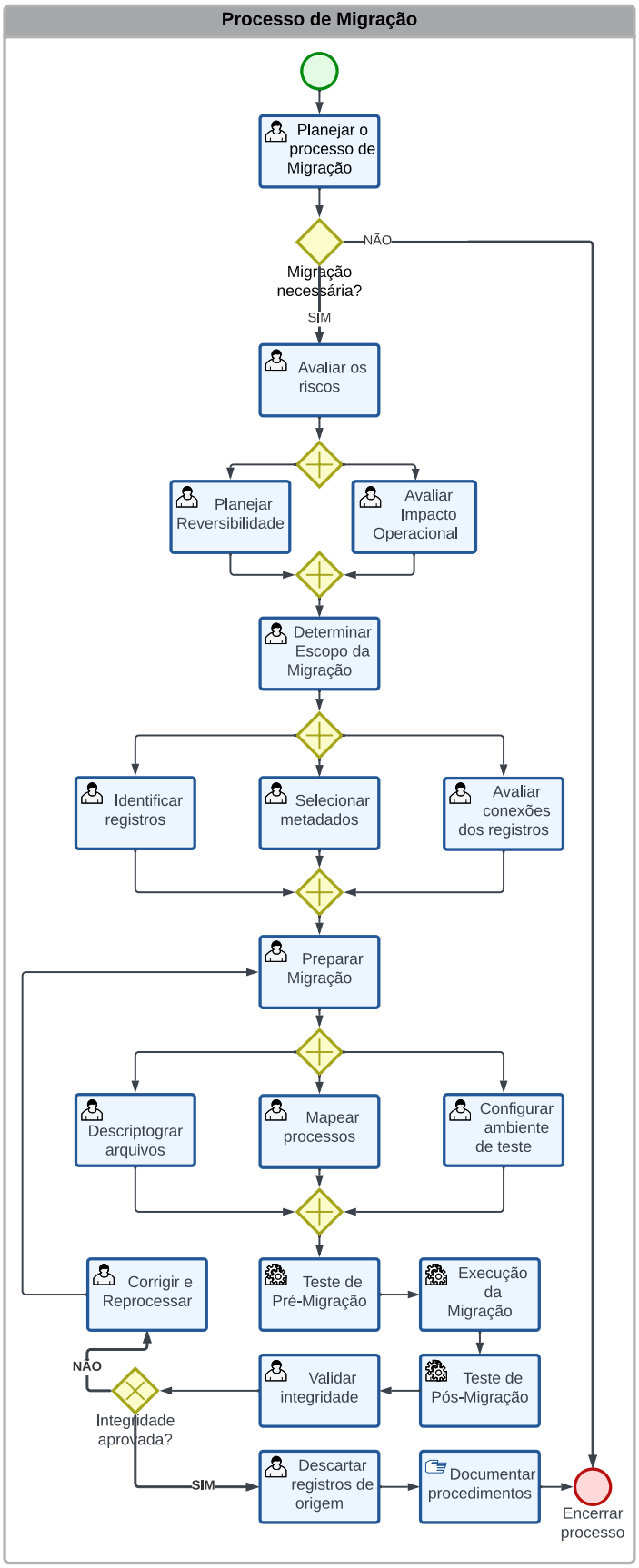
- Definir padrões de *hardware* e *software* que sejam compatíveis com a infraestrutura institucional; e
- Selecionar formatos de armazenamento adequados ao *hardware* e *software* adotados (GRÁCIO, 2012).

Embora não seja considerado uma técnica de preservação quando aplicado isoladamente, o refrescamento periódico de suporte configura uma importante atividade para garantir o acesso aos objetos digitais, juntamente com a verificação da integridade do suporte, conforme previsto na etapa de *armazenamento*, para que seja aplicado quando o suporte estiver se tornando obsoleto ou for identificado sinais de degradação (FERREIRA, 2006, SANTOS; FLORES, 2015b). O refrescamento atua na preservação do nível físico do objeto, desta forma, poderá reduzir os efeitos da obsolescência em nível de suporte, bem como os impactos da degradação do mesmo (SANTOS; FLORES, 2015b).

Também essencial para o armazenamento, os processos de migração devem ocorrer periodicamente. Recomenda-se a migração parcial de *hardware* na instituição, que envolve o revezamento de computadores pessoais com máquinas novas. Esse processo permite que os equipamentos mais novos sejam utilizados nas áreas que demandam maior desempenho e atualização tecnológica constante, como nas áreas de laboratórios, radiologia e pronto-socorro, enquanto as máquinas substituídas, ainda em bom estado, podem ser redistribuídas para setores com necessidades menos intensivas, como áreas administrativas. Dessa forma, a instituição garante uma alocação mais eficiente de recursos tecnológicos, prolonga a vida útil dos equipamentos e reduz custos.

À medida que novas versões dos softwares são disponibilizadas pelas equipes de desenvolvimento dos sistemas utilizados na instituição, é necessário realizar a atualização de versões, garantindo a compatibilidade dos dados tanto na base local quanto com o DATASUS. Já a migração de formato (conversão) deve ser realizada conforme os formatos mais adequados à preservação discutidos na *Gestão de Riscos*. Esses processos de migração devem ser aplicados continuamente e, por isso, requerem monitoramento constante para garantir sua correta execução e eficácia. A Figura 5.4 ilustra o fluxo das etapas para implementação do processo de migração e foi baseado nas recomendações do Governo de Queensland (2021).

Figura 5.4 – Fluxo do processo de Migração



Para representar as atividades a serem realizadas, foi utilizado o *Business Process Model and Notation* (BPMN), uma notação padrão voltada para a modelagem de processos. Essa notação facilita a representação visual das atividades, interações e decisões necessárias e oferece uma visão clara e compreensível de cada passo do fluxo de trabalho.

Na primeira etapa, é feito o planejamento para identificar se a migração é realmente necessária, considerando a substituição ou descomissionamento de sistemas. Considerando a necessidade de migração, ocorre a avaliação de riscos, que identifica possíveis ameaças à integridade dos dados. Uma estratégia de reversão é então definida para permitir o retorno ao sistema original, caso a migração apresente problemas. Também é definido o impacto operacional, que considera como os processos e funcionários serão afetados.

Na Determinação do escopo da migração, são identificados quais registros precisam ser migrados e quais podem ser destruídos antes do processo. É feita uma seleção de metadados que devem ser mantidos para garantir a continuidade dos registros após a migração. Além disso, é essencial garantir a manutenção de conexões entre os registros e suas informações contextuais e estruturais, para que nenhum dado perca seu significado ou usabilidade. Com o escopo definido, inicia-se a preparação técnica. Isso inclui a descriptografia de arquivos que necessitam de acesso durante a migração, além de um mapeamento de processos para entender melhor o sistema de origem e as relações entre registros e processos de negócios. Também é configurado um ambiente de teste que simula o processo de migração, permitindo verificar o sucesso do processo antes de realizá-lo no ambiente real.

Antes de realizar a migração definitiva, é necessário testar a migração no ambiente de teste. Isso inclui a verificação de integridade dos registros e metadados, assegurando que todas as informações serão preservadas. A análise de riscos é revisada para confirmar que os riscos foram minimizados, e qualquer problema identificado durante o teste deve ser resolvido antes da migração final. Os resultados desses testes devem ser documentados. A migração é então realizada e durante o processo deve ocorrer um monitoramento em tempo real para detectar e corrigir quaisquer problemas que possam surgir.

Após a migração, deve ser realizado um novo ciclo de testes, agora no sistema real. Estes testes de pós-migração verificam a integridade e a confiabilidade dos registros

no novo ambiente. Uma verificação de qualidade é feita para garantir que o novo sistema está gerenciando os registros corretamente e que nenhum dado foi alterado de forma não autorizada. Somente após a confirmação de que a migração foi bem-sucedida, deve ser autorizada a destruição dos registros de origem. Isso deve ocorrer apenas quando todos os registros migrados forem verificados e aprovados no novo sistema. A etapa final é a documentação completa do processo de migração, incluindo metadados sobre a migração e destruição, detalhes do processo, datas, participantes e todas as aprovações necessárias.

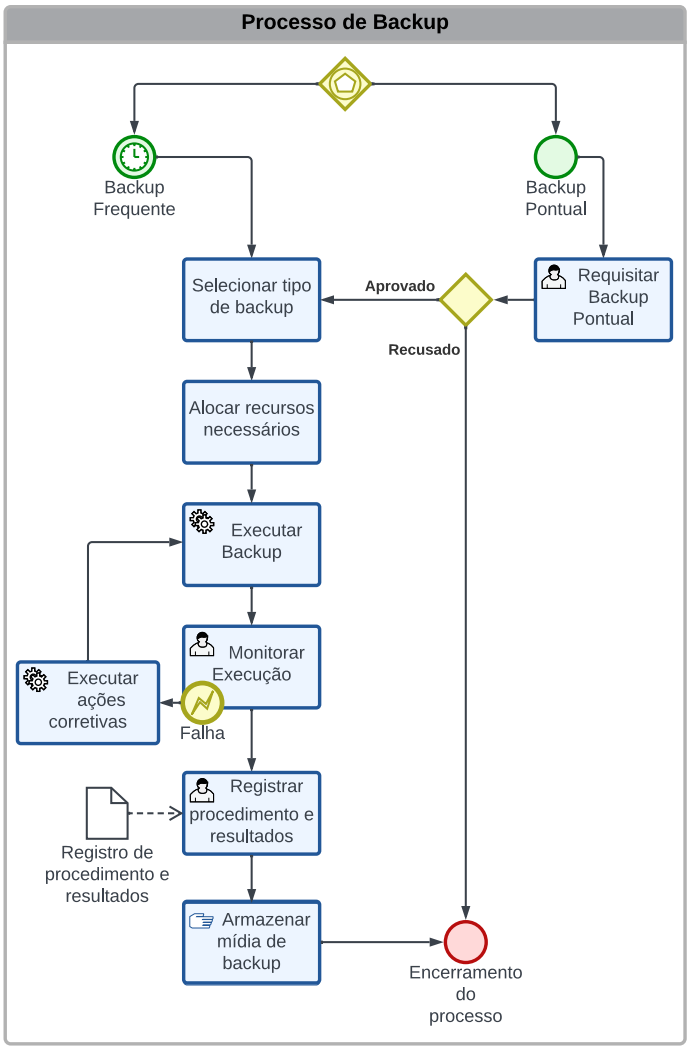
Embora não seja definido como uma estratégia de preservação, a manutenção de cópias redundantes de conteúdo digital é uma medida de segurança essencial para a Unidade de Saúde São Lucas. O *backup* tem um papel específico de garantir a recuperação em caso de perda ou corrupção de dados, e é uma prática que garante a integridade dos dados a curto e médio prazo. Além de realizar os procedimentos de *backup* é igualmente importante testar periodicamente a mídia de backup para garantir que os dados permaneçam legíveis e não foram alterados, assim como testar os procedimentos de restauração, verificando se o *hardware*, *software* e fornecedores envolvidos estão funcionando conforme esperado. O *backup* deve ser abrangente, que inclui o sistema operacional, os *softwares* e todos os objetos digitais do sistema.

O processo de geração de arquivos de *backup* deve ser realizado e armazenado obrigatoriamente em uma área ou dispositivo dedicado, de alta capacidade, instalado localmente e disponível *online* na rede interna da instituição, com acesso restrito. Além disso, é recomendável que o *backup* seja replicado em dispositivos de armazenamento fora da rede de computadores (*offline*) e fora da rede elétrica, como em fitas magnéticas de padrão aberto *Linear Tape-Open* (LTO) (ARQUIVO PÚBLICO DE SÃO PAULO, 2022).

O processo de *backup* representado na Figura 5.5 se inicia com dois possíveis cenários: o *backup* frequente, que é configurado para ocorrer de forma automatizada em intervalos predefinidos, e o *backup* pontual, solicitado manualmente em casos específicos de necessidade imediata e que necessita de aprovação prévia dos responsáveis pelo processo. O próximo passo é a seleção do tipo de *backup* (completo, incremental ou diferencial) e a alocação dos recursos necessários para sua execução. Essa etapa envolve identificar os dispositivos de armazenamento e garantir

que o ambiente esteja preparado para realizar a cópia de segurança de forma eficiente.

Figura 5.5 – Fluxo do processo de Backup



O *backup* é então executado por ferramentas específicas, podendo ser um processo manual ou automatizado. Durante a execução, deve ser realizado um monitoramento constante para identificar possíveis falhas. Caso uma falha seja detectada, o fluxo do processo redireciona para a execução de ações corretivas, como análise de logs, ajustes na configuração ou uma nova tentativa de *backup*. Após a execução, os resultados são registrados, documentando o sucesso ou os problemas encontrados no processo. O sistema de cópias de segurança deve dispor de uma trilha de auditoria para permitir a recuperação dos materiais digitais gerados no intervalo entre falhas.

Os *backups* bem-sucedidos são armazenados nas mídias apropriadas e então o processo é encerrado.

Esse plano reforçou a importância de aplicar diversas estratégias complementares para mitigar riscos associados à manutenção de objetos digitais. Cada tipo de ameaça exige uma abordagem específica, o que torna evidente a necessidade de diversificação das técnicas e metodologias empregadas. A Tabela 5.11 indica como os elementos apresentados se complementam, formando uma camada abrangente de proteção.

Tabela 5.11 – Ameaças e vulnerabilidades da preservação

Ameaças e Vulnerabilidades			**Técnicas/estratégias**					
			Redundância	Migração	Refrescamento	Diversidade	Metadados	Auditoria
Vulnerabilidades	Dados	Falhas de mídia	R	r	r	-	R	R
		Obsolescência de mídia	-	r	-	-	R	R
	Infraestrutura	Falhas de *hardware*	-	r	r	r	-	R
		Obsolescência de *hardware*	-	r	r	r	-	R
		Falhas de comunicação	-	-	r	r	-	R
		Falhas de serviço de rede	-	-	r	r	-	R
	Processo	Falhas de *software*	-	r	r	r	-	R
		Obsolescência de *software*	-	r	r	r	-	R
Ameaças	Desastres	Desastres naturais	R	-	-	r	-	-
		Erros operacionais humanos	R	-	-	r	R	R
	Ataques	Ataques internos	R	-	-	r	R	R
		Ataques externos	R	-	-	r	R	R
	Gestão	Falhas econômicas	-	-	-	r	-	R
		Falhas organizacionais	-	-	-	r	-	R
	Legislação	Mudanças legislativas	-	-	-	r	r	-

Legenda: r = reduz o risco da ameaça/vulnerabilidade; R = requerido para recuperação; - = não se aplica.

Fonte: adaptado de BARATEIRO et al. (2010)

5.3.6 Acesso

O principal objetivo da preservação digital é assegurar que os dados permaneçam acessíveis e utilizáveis a longo prazo. Na etapa de *acesso*, devem ser aplicados mecanismos de segurança e controles de acesso para garantir que as informações armazenadas sejam consultadas apenas por indivíduos autorizados, em conformidade com as normas de proteção de dados e as exigências legais aplicáveis (LIRA; SIEBRA, 2022).

No contexto da preservação de registros médicos, como prontuários de pacientes, esse aspecto deve considerar as obrigações legais e éticas relacionadas ao sigilo das informações pessoais sensíveis. No Código de Ética Médica, artigo 73, é vedado ao médico revelar publicamente ou a terceiros qualquer informação que tenha obtido em razão de sua atividade profissional, salvo em situações específicas. Esse sigilo pode ser quebrado apenas com a autorização expressa do paciente (titular dos dados), por determinação judicial ou para a defesa do próprio médico (CFM, 2019). Essa proibição deve se estender a qualquer profissional da Unidade de Saúde São Lucas que tenha acesso à dados dos pacientes.

No caso do PEC, caso o profissional de saúde deseje visualizar o prontuário sem a presença do paciente, é obrigatório o registro de uma justificativa para acessar as informações clínicas fora do contexto do atendimento presencial. Essa justificativa é armazenada no banco de dados e pode ser utilizada posteriormente em auditorias, assegurando que o acesso às informações sensíveis respeite as normas de sigilo e garantindo, assim, a privacidade e segurança dos dados clínicos do cidadão.

O PEC, PACS e SISREG adotam um sistema de perfis de acesso, no qual cada profissional possui um controle de permissões e auditoria vinculado ao seu login pessoal, que é feito pelo CPF. Dessa forma, o sistema exige que os profissionais utilizem exclusivamente seu próprio login e senha, estabelecendo uma camada adicional de responsabilidade e rastreabilidade no acesso aos dados dos pacientes.

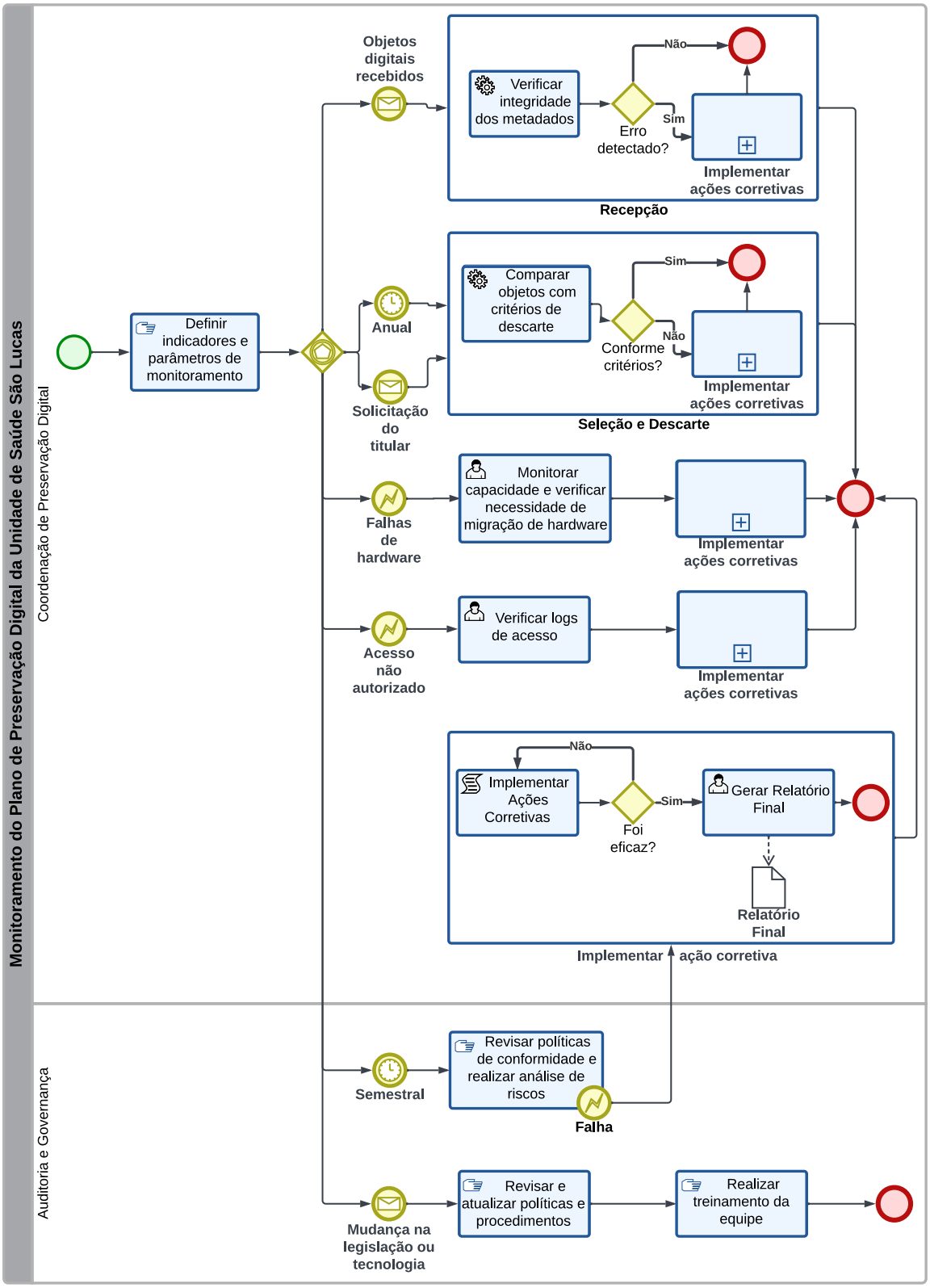
De forma complementar à essas ações, o acesso físico a lugares onde os computadores são mantidos deve ser restringido. Por isso, a equipe responsável pelo acesso aos dados deve estar adequadamente treinada e ciente das responsabilidades legais e éticas associadas, de forma a prevenir o uso inadequado das informações.

5.4 Monitoramento e Revisão

Para a preservação digital, o processo de monitoramento é a atividade de acompanhamento e adequação constante dos processos que abrangem o modelo às necessidades dos usuários e da instituição (GRÁCIO, 2012). No diagrama apresentado na Figura 5.6, foram considerados dois atores principais (representados por *lanes* no diagrama): a Coordenação de Preservação Digital e a Auditoria de Governança, que têm responsabilidades distintas e complementares ao longo do ciclo de monitoramento do plano de preservação digital. A Coordenação de Preservação Digital é a equipe responsável por garantir a integridade e a continuidade dos objetos digitais. O ponto de partida é a definição dos critérios para acompanhamento do plano de preservação digital. A partir dessa atividade, de forma paralela, diferentes eventos podem ser o gatilho para as atividades posteriores.

Quando os objetos digitais são recebidos (evento de mensagem), há uma primeira etapa de verificação da integridade dos metadados. A Coordenação é responsável por avaliar se há algum erro nos metadados e, caso exista, implementar ações corretivas antes de dar continuidade ao fluxo. Esse processo é essencial para garantir que os objetos estejam em conformidade desde o início. De forma periódica (anualmente, conforme mostrado no diagrama) ou por solicitação do titular, ocorre a análise dos objetos digitais com base em critérios de descarte preestabelecidos. Essa análise é importante para evitar a sobrecarga dos sistemas. A Coordenação deve implementar ações corretivas caso o objeto precise ser descartado.

Figura 5.6 – Monitoramento do Plano de Preservação Digital



Fonte: elaborado pela autora

A Coordenação também se preocupa com a manutenção do *hardware* utilizado para a preservação digital, monitorando a capacidade de armazenamento e verificando se há necessidade de migração para um novo *hardware*. Esse monitoramento inclui sinais de obsolescência ou falhas temporárias que podem gerar problemas para a preservação. Além disso, em situações de acesso não autorizado, os *logs* de acesso devem ser verificados para garantir a segurança dos objetos digitais. Caso algum problema seja identificado em qualquer uma das etapas descritas acima, ações corretivas são implementadas de acordo com o plano de preservação digital da instituição. Posteriormente, a eficácia dessas ações é verificada. Se as ações forem eficazes, um relatório final é gerado para documentar o processo e os resultados obtidos.

A Auditoria de Governança atua como um segundo ator importante, focado na revisão e na atualização de políticas e na garantia de que as práticas de preservação estejam alinhadas às normas vigentes e objetivos da instituição. De forma semestral, as políticas de conformidade são revisadas, e realiza-se uma análise de riscos para identificar pontos fracos e potenciais melhorias. Em caso de falhas encontradas, ações corretivas são tomadas para adequar os processos. No caso de mudanças legislativas ou tecnológicas, a Auditoria de Governança tem como responsabilidade revisar e atualizar as políticas e procedimentos vigentes. Após essas atualizações, é necessário treinar a equipe, para que todos estejam cientes das novas práticas e das mudanças implementadas.

Os gatilhos que impulsionam as atividades no ciclo de monitoramento podem ser monitorados por meio de ferramentas automatizadas, que facilitam a detecção e resposta rápida a determinados esses eventos. A utilização do *Fixity Pro*, por exemplo, permite verificações regulares programadas da integridade dos arquivos e gera relatórios detalhados sobre o status dos arquivos, destacando principalmente as alterações detectadas durante as verificações (COPTR, 2021). Da mesma forma, a utilização de sistemas específicos, como o Archivematica, pode ser configurado para validar a integridade dos metadados dos objetos digitais no momento da recepção, estabelecendo um ponto inicial de conformidade.

No que se refere ao monitoramento de armazenamento e *hardware*, ferramentas como Zabbix[47] e Nagios[48], ambas de código aberto, pode ser configuradas para emitir gatilhos automáticos quando o armazenamento atinge limites críticos ou quando sinais de obsolescência de hardware, como aumento na taxa de erros de leitura e gravação, desempenho degradado, temperaturas elevadas ou mensagens de diagnóstico, são detectados. Em um ambiente hospitalar, esse monitoramento é um processo não apenas essencial, mas, em muitos casos, vital para salvar vidas. Uma solução de monitoramento confiável assegura que o fluxo de dados necessário ocorra com máxima rapidez e precisão, permitindo que os profissionais de saúde se concentrem nos procedimentos médicos necessários (ZABBIX, [s.d.]).

O Zabbix é um *software* de monitoramento que cobre a a integridade de redes, servidores, máquinas virtuais, aplicações, serviços, bancos de dados, websites e ambientes na nuvem, entre outros. Ele possui um mecanismo flexível de notificação que permite configurar alertas personalizados via *e-mail* para praticamente qualquer tipo de evento. Além disso, o Zabbix oferece recursos avançados de relatórios e visualização de dados gratuitamente (ZABBIX, [s.d.]). De forma semelhante, o Nagios também permite a configuração de alertas automáticos para acompanhar o desempenho de servidores, aplicações e dispositivos, mas parte dos serviços são oferecidos sob uma licença comercial.

Conforme as atividades e ações de preservação digital são realizadas, pode ser necessário adaptações nos formatos dos dados, nos suportes, na infraestrutura tecnológica, nos métodos de distribuição e acesso, no encaminhamento do objeto digital para preservação, entre outros aspectos (GRÁCIO, 2012). A Figura 5.7 ilustra esse processo.

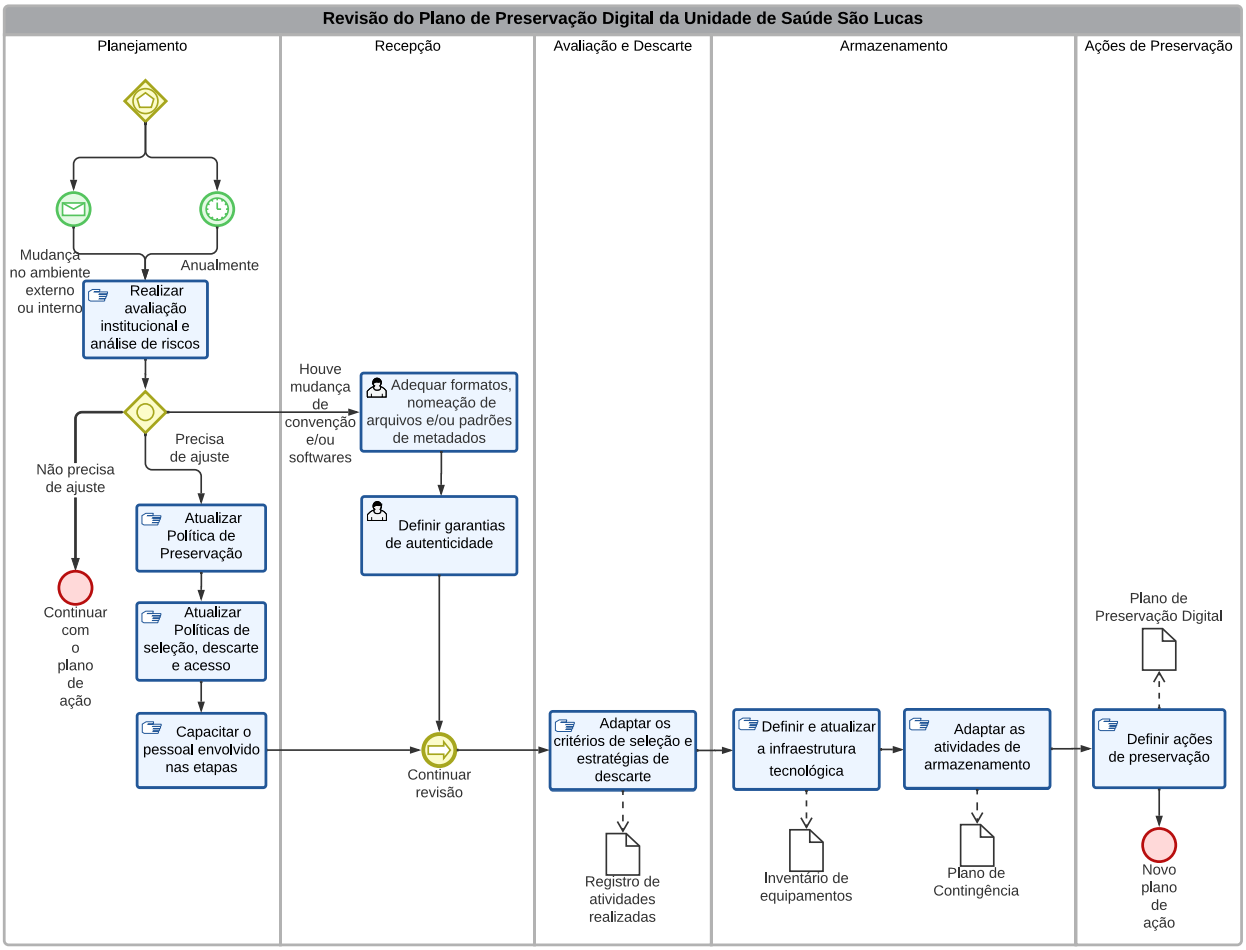
---

[47] Disponível em https://www.zabbix.com/

[48] Disponível em https://www.nagios.org/

O diagrama que descreve as atividades de revisão também foi elaborado utilizando o BPMN, conforme Figura 5.7Figura 5.7. Esse processo envolve as etapas de *planejamento*, *recepção*, *seleção e descarte*, *armazenamento* e *ações de preservação* do Plano de preservação digital da Unidade de Saúde São Lucas, divididas em *lanes* específicas. O *planejamento* é a etapa que define o início de como o processo de revisão do plano será conduzido.

Figura 5.7 – Revisão do Plano de Preservação Digital



Fonte: elaborado pela autora

Nessa etapa, consideramos eventos externos e internos que podem motivar uma reavaliação das políticas e práticas de preservação, como mudanças legislativas, tecnológicas, ou novas diretrizes e estrutura institucionais. Essas mudanças também devem estar sendo monitoradas. Além disso, há também as revisões periódicas do plano, que ocorrem anualmente para que a preservação digital na instituição esteja sempre atualizada e alinhada com as boas práticas. Essa revisão é alinhada ainda com os requisitos da LGPD.

A partir de algum desses eventos, a realização de uma análise institucional e de riscos, auxilia a identificar possíveis falhas ou pontos de melhoria a serem aprimorados no plano de preservação digital. Se forem identificadas necessidades de melhoria, três ações são implementadas: 1. Atualização dos principais pontos da Política de Preservação; 2. Atualização das Políticas de Seleção, Descarte e Acesso; e 3. Capacitação do Pessoal Envolvido nas etapas de preservação. É importante ressaltar que essas atualizações devem ocorrer de acordo a necessidade da instituição e os problemas identificados.

Na *recepção*, o diagrama apresenta a etapa em que se verifica se houve alguma mudança na convenção adotada ou nos *softwares* utilizados, o que pode exigir ajustes nos formatos dos objetos digitais ou padrões adotados. Essas mudanças também implicam na atualização das medidas e critérios para garantir a autenticidade dos dados. Caso mudanças sejam necessárias, adaptações são feitas nos critérios de seleção e estratégias de descarte, que são então registrados.

Na etapa de *armazenamento*, o objetivo é assegurar que a capacidade de armazenamento e as tecnologias empregadas estão adequadas às mudanças realizadas no ambiente. As atividades incluem a atualização da infraestrutura de TI e das atividades de armazenamento, que geram novos documentos de inventario de equipamentos e o plano de contingência, respectivamente. Por fim, nas *ações de preservação*, são definidas as estratégias operacionais de preservação digital que estejam de acordo com as novas políticas elaboradas. Esse ponto marca a criação de um novo plano de ação, que incorpora todas as mudanças e melhorias identificadas ao longo do processo.

6 Avaliação

Para uma avaliação estruturada do modelo dessa tese, o *Quantitative Evaluation Framework* (QEF) foi utilizado. O modelo de avaliação permite uma análise estruturada e quantitativa dos elementos do modelo, considerando tanto os aspectos técnicos quanto operacionais. ARELLANO (2008) definiu os seguintes seis critérios fundamentais para a preservação digital:

- **Confiabilidade**: Este critério inclui os requisitos técnicos e gerenciais que garantem a integridade dos formatos, a segurança e a permanência do armazenamento dos dados ao longo do tempo. A confiabilidade se manifesta em requisitos como a autenticação de dados, garantindo que não foram alterados, e a realização de verificações de vírus.
- **Responsabilidade Política**: Esse critério aponta para a responsabilidade da instituição em manter e preservar os acervos digitais. Ela deve ter uma política clara sobre quem tem acesso aos dados e em quais condições. Neste caso, a Governança envolve a criação de políticas e planos que asseguram que a preservação esteja alinhada com a visão institucional.
- **Sustentabilidade Econômica**: Este critério foca na viabilidade financeira de longo prazo da implementação do plano de preservação digital, exigindo ações que garantam que o serviço seja sustentável ao longo do tempo. Os requisitos relacionados podem incluir a alocação de recursos humanos e financeiros para suportar a preservação digital de forma contínua.
- **Inclusão em Repositórios Digitais**: A inclusão em repositórios digitais promove a validação e o reconhecimento científico dos dados e serviços preservados. A interoperabilidade e a eficácia são essenciais para garantir que os dados possam ser trocados e usados em diversos sistemas e ambientes.
- **Transparência**: A transparência exige que as especificações técnicas estejam documentadas, permitindo auditoria e certificação do conteúdo. Esse critério é abordado em requisitos como a documentação das ações realizadas e o registro de quem teve acesso aos dados, assegurando um histórico completo para auditorias futuras.

- **Acessibilidade de Longo Prazo**: Este critério visa garantir que os dados permaneçam acessíveis no longo prazo. Inclui requisitos de manutenção técnica, interoperabilidade, desempenho e conectividade com outros objetos e serviços. A fixidez dos dados e a atualização dos sistemas de preservação também são considerados para assegurar o acesso futuro.

Esses critérios foram adaptados para definição dos requisitos do QEF e baseados em outras recomendações gerais de boas práticas para preservação digital e nas especificidades do ambiente, e então divididos em 3 dimensões: Domínio Técnico, Domínio Organizacional e Domínio de Recursos. O Domínio Técnico abrange os aspectos relacionados à tecnologia e às ferramentas de preservação digital, são eles:

Tabela 6.1 – Características do Domínio Técnico

Id	*Aspecto Domínio Técnico*
A1	Autenticidade
A2	Disponibilidade
A3	Interoperabilidade
A4	Eficácia

O critério *A1 – Autenticidade* envolve a inclusão de metadados, controles de integridade e verificações regulares contra ameaças que possam comprometer os arquivos. Possui três requisitos: *TA01 – Metadados são mantidos e corretamente associados ao conteúdo digital*, envolve a capacidade de armazenar e manter metadados precisos e completos que estejam permanentemente vinculados ao objeto digital; o requisito *TA02 – Garantia de que os dados não foram alterados ou corrompidos* trata da implementação de mecanismos e procedimentos para garantir que os dados permaneçam inalterados desde sua criação ou última modificação autorizada. Isso pode incluir o uso de métodos de verificação que identificam qualquer alteração não autorizada nos dados; e o *TA03 – Realizar verificações de vírus* foca na proteção dos arquivos contra ameaças externas que possam comprometer a integridade dos dados, e envolve a implementação de verificações regulares de vírus, usando ferramentas e *softwares* de segurança para identificar e limpar ameaças identificadas.

A *Disponibilidade (*critério *A2*) busca definir critérios para evitar a perda de informações críticas e garantir a continuidade do acesso. O requisito *TD01 – Implementação de planos de continuidade que garantam o acesso aos dados em caso de desastre* verifica se o plano inclui estratégias e medidas de contingência para responder a situações que possam comprometer a disponibilidade dos dados. Este plano de emergência deve minimizar o impacto de eventos como falhas de *hardware*, desastres naturais, ataques cibernéticos ou qualquer interrupção que ameace o acesso ao conteúdo digital.

O *TD02 – Seleção e atualização de hardwares e softwares que ofereçam as melhores expectativas de garantia de acesso a longo prazo* refere-se ao uso de tecnologias que sustentem a preservação digital de forma duradoura. Esse requisito envolve critérios para escolher e atualizar os *hardwares* e *softwares* usados no armazenamento e acesso aos dados, considerando fatores como obsolescência tecnológica e compatibilidade futura. Por último, o requisito *TD03 – Registros de quem realizou quais ações nos arquivos* diz respeito à implementação de sistemas de auditoria e monitoramento que documentam todas as ações realizadas nos arquivos preservados.

O critério *A3 – Interoperabilidade* envolve a escolha de formatos adequados, a conversão de arquivos legados e a integração entre sistemas. Os requisitos são: *TI01 – Utilização de formatos que sejam compatíveis com normas e padrões nacionais e internacionais de preservação digital* garante que os formatos de arquivo selecionados estejam em conformidade com os padrões amplamente aceitos na área de preservação digital, mais propensos a receber suporte contínuo e a serem legíveis por sistemas futuros; *TI02 – Implementação de processos para conversão de arquivos de formatos legados para formatos modernos e interoperáveis* aborda a necessidade de converter arquivos em formatos antigos para formatos mais atuais e avalia a eficácia dos processos implementados para realizar essas conversões de forma segura; e *TI03 – Integração e troca de dados entre os diferentes sistemas* mede a capacidade da instituição de integrar e compartilhar dados de forma eficiente entre seus sistemas internos e externos.

O último critério deste domínio, *A4 – Eficácia*, em um plano de preservação digital avalia se as ferramentas e sistemas utilizados para a preservação de dados estão operando de forma otimizada, segura e eficiente. O requisito *TE01 – Implementação de ferramentas de monitoramento para garantir que as ferramentas de preservação*

estejam funcionando de forma otimizada avalia a adoção de ferramentas de monitoramento que acompanham as diversas atividades e processos de preservação digital. O *TE02 – Manter um plano regular de atualizações para garantir que as ferramentas estejam sempre seguras e funcionais* destaca a importância de manter um cronograma de atualizações para todas as ferramentas e sistemas usados no processo de preservação digital. Esse plano de atualizações inclui sistemas de armazenamento, *softwares* de monitoramento e outras aplicações críticas para o gerenciamento de dados.

A dimensão do domínio organizacional refere-se às políticas, práticas e estratégias que sustentam a preservação digital ao longo do tempo. A Sustentabilidade e Governança são aspectos que suportam as ações de preservação e devem ser avaliadas de acordo com os objetivos da instituição.

Tabela 6.2 – Características do Domínio Organizacional

Id	*Aspecto Domínio Organizacional*
A6	Sustentabilidade
A7	Governança

A *A6 – Sustentabilidade* em um plano de preservação digital é fundamental para garantir que as atividades de preservação possam continuar a longo prazo, apesar das mudanças tecnológicas e de eventuais desafios financeiros ou organizacionais. O requisito *OS01 – Definição e documentação de etapas futuras para a preservação digital* prevê o planejamento e à criação de uma estratégia para garantir a continuidade das atividades de preservação digital. Esse requisito envolve a documentação das ações futuras que serão necessárias para adaptar o plano de preservação ao longo do tempo, incluindo a definição de metas de curto, médio e longo prazo.

O *OS02 - Identificação e mitigação de riscos que possam comprometer a sustentabilidade do plano de preservação a longo prazo* diz respeito à avaliação e mitigação de ameaças que possam afetar a continuidade do plano de preservação digital. Esse requisito exige a identificação de potenciais riscos, como falhas de *hardware*, mudanças no financiamento, ou desafios organizacionais, e a criação de estratégias para mitigar esses riscos. Por fim, o requisito *OS03 - Implementação de*

sistemas que sejam flexíveis para se adaptarem às mudanças tecnológicas avalia a capacidade dos sistemas de preservação digital de se ajustarem a novas tecnologias e padrões. Esse requisito inclui a compatibilidade com novos formatos de arquivo, a capacidade de migração de dados para tecnologias mais recentes e a modularidade da arquitetura do sistema, o que permite uma adaptação mais ágil a mudanças futuras.

O critério de Governança considera a estrutura organizacional e as políticas que orientam as atividades de preservação. O *requisito OG01 - Elaboração de uma Política de Preservação Digital* avalia a existência e a qualidade de uma política formal que guie todas as atividades de preservação digital na instituição. Para o cumprimento total desse requisito, a política deve ser abrangente, definindo princípios, responsabilidades e diretrizes para a preservação a longo prazo dos dados.

O *requisito OG02 - Definição clara de políticas de acesso que determinem quem pode acessar os dados e em que condições* garante que a instituição tenha políticas de acesso documentadas, estabelecendo quem tem permissão para acessar os dados preservados e em quais condições. O controle de acesso claro e documentado ajuda a prevenir acessos não autorizados, especialmente importante para dados pessoais sensíveis. Já o requisito *OG03 - Garantir que as metas de preservação digital estejam alinhadas com a visão, missão e normas da instituição*, destaca a importância de que as metas e atividades de preservação digital sejam coerentes com os valores e objetivos estratégicos da instituição. O alinhamento torna a preservação digital não apenas um aspecto técnico, mas também um componente essencial da estratégia organizacional.

Por último, o Domínio de Recursos aborda os aspectos financeiros, a análise de custos e os recursos humanos relacionados à aplicação do modelo de preservação digital. A Relação custo-benefício aborda a viabilidade econômica e a eficiência dos gastos com o modelo. E o Recursos Humanos enfoca a importância de ter uma equipe bem treinada, capacitada e motivada, essencial para a implementação e manutenção eficazes das práticas de preservação digital.

Tabela 6.3 – Características do Domínio de Recursos

Id	*Aspecto Domínio de Recursos*
A8	Relação custo-benefício
A9	Recursos Humanos

O critério *A8 – Relação Custo-Benefício* considera que a preservação digital envolve custos contínuos em várias etapas do ciclo de vida dos dados, desde a captura inicial até a preservação a longo prazo. O requisito *RCB01 - Avaliar todos os custos ao longo do ciclo de vida dos arquivos, desde a captura até a preservação a longo prazo* exige uma análise completa dos custos associados à preservação digital. Essa avaliação deve abranger todas as fases do ciclo de vida dos arquivos digitais, incluindo captura, armazenamento, migração, monitoramento e preservação. A análise detalhada de custos permite que a instituição identifique as áreas que demandam mais recursos, otimize o uso desses recursos e planeje financeiramente para garantir que os dados possam ser mantidos e acessados no futuro, sem comprometer a estabilidade financeira da instituição.

O critério *A9 – Recursos Humanos* é outro aspecto crucial em um plano de preservação digital, pois as pessoas envolvidas são responsáveis pela execução, manutenção e evolução das atividades de preservação. O requisito *RH01 – Definição clara de papéis e responsabilidades relacionadas às ações de preservação digital* avalia a documentação e a clareza das funções de cada membro da equipe, de forma que todos os envolvidos saibam exatamente quais são suas responsabilidades em cada etapa do processo de preservação digital. O requisito *RH02 – Desenvolvimento de programas de treinamento contínuo para manter a equipe atualizada sobre as melhores práticas de preservação digital* avalia se o plano abrange treinamentos regulares para a equipe, de forma que os profissionais envolvidos se mantenham atualizados sobre as melhores práticas, tecnologias e normas do setor.

As dimensões que compõem a qualidade do modelo têm diferentes níveis de importância dependendo do contexto. Essa qualidade é medida então com base nesse contexto, no seu enquadramento e finalidade. Para o cálculo de qualidade, os aspectos recebem o mesmo peso e a média desses aspectos é usada para calcular a qualidade (ESCUDEIRO; BIDARRA, 2008). Devemos definir as importâncias relativas de cada aspecto para cada uma das dimensões:

$$\sum_{\mathrm{n}} (\mathrm{p_n} \times fator_{\mathrm{n}})$$

Onde,

n é o número de aspectos relevantes para a dimensão

p_n é o peso do aspecto n na dimensão, sendo $p_n = 1$, e

$fator_n$ é a avaliação do aspecto n.

Cada aspecto é avaliado da seguinte forma:

$$\frac{1}{\sum_m pr_m} \times \sum_m (pr_m \times pc_m)$$

Onde,

m é o número de critérios relevantes para o aspecto em análise

pr_m é o peso do critério m, e

pc_m é a porcentagem de cumprimento do critério m

A qualidade de um sistema é uma medida, da distância entre o sistema ideal, e o sistema produzido. O desvio global é então dado por:

$$D = \sqrt{\sum_j \left(1 - \frac{Dim_j}{100}\right)^2}$$

Por fim, a qualidade do sistema é inversamente proporcional à D e é dada por:

$$Q = 1 - \frac{D}{\sqrt{n}}$$

A avaliação foi iniciada classificando os requisitos dos critérios definidos, para cada domínio, com um peso entre 0 e 10, de acordo com a sua relevância para cada uma das dimensões. A classificação é determinada por: 10 – Fundamental; 8 – Muito importante, 6 – Importante; 4 – Necessário; 2 – Opcional; 0 – Irrelevante (ESCUDEIRO; BIDARRA, 2008). A classificação e avaliação com todos os cálculos necessários para obter a qualidade do modelo estão disponíveis no *software* Microsoft Excel (ESCUDEIRO; BIDARRA, 2008).

Para avaliar o cumprimento do plano de preservação digital, foi elaborada uma tabela que atribui uma porcentagem de cumprimento aos testes realizados em cada aspecto avaliado. Essa tabela permite quantificar o grau de implementação dos requisitos e identificar áreas onde o plano atende plenamente às exigências, bem como aquelas que ainda precisam de melhorias. Cada requisito é analisado em relação ao seu cumprimento, e a partir dessa análise, é possível determinar a conformidade geral do plano com as melhores práticas de preservação digital. A Figura 6.1 ilustra o modelo QEF devidamente preenchido.

Figura 6.1 – Avaliação de qualidade do modelo

q	D	x	Domínio	Qj	Wij (Peso do Critério j no Domínio i) [0,1]	Critério (j)	rwjk (peso do requisito k no Critério j) {2, 4, 6, 8, 10}	Requisito	wfk % cumprimento do requisito k [0, 100]
72%	0,724	77,3	Técnico	83,33333	0,273	Autenticidade	10	TA01. Metadados são mantidos e corretamente associados ao conteúdo digital.	75
							10	TA02. Garantia de que os dados não foram alterados ou corrompidos.	75
							10	TA03. Realizar verificações de vírus.	100
				82,14286	0,273	Disponibilidade	10	TD01. Implementação de planos de continuidade que garantam o acesso aos dados em caso de desastre.	50
							10	TD02. Seleção e atualização de hardwares e softwares que ofereçam as melhores expectativas de garantia de acesso à longo prazo.	100
							8	TD03. Registros de quem realizou quais ações nos arquivos.	100
				67,85714	0,273	Interoperabilidade	10	TI01. Utilização de formatos que sejam compatíveis com normas e padrões nacionais e internacionais de preservação digital.	100
							8	TI02. Implementação de processos para conversão de arquivos de formatos legados para formatos modernos e interoperáveis.	50
							10	TI03. Integração e troca de dados entre os diferentes sistemas.	50
				75	0,182	Eficácia	10	TE01. Implementação de ferramentas de monitoramento para garantir que as ferramentas de preservação estejam funcionando de forma otimizada.	50
							10	TE02. Manter um plano regular de atualizações para garantir que as ferramentas estejam sempre seguras e funcionais.	100
		58,4	Organizacional	44,64286	0,429	Sustentabilidade	10	OS01. Definição e documentação de etapas futuras para a preservação digital.	75
							10	OS02. Identificação e mitigação de riscos que possam comprometer a sustentabilidade do plano de preservação a longo prazo.	50
							8	OS03. Implementação de sistemas que sejam flexíveis para se adaptarem às mudanças tecnológicas.	0
				68,75	0,571	Governança	10	OG01. Elaboração de uma Política de Preservação Digital.	50
							10	OG02. Definição clara de políticas de acesso que determinem quem pode acessar os dados e em que condições.	100
							10	OG03. Garantir que as metas de preservação digital estejam alinhadas com a visão, missão e normas da instituição.	100
							10	OG04. Documentação de procedimentos operacionais padrão para todas as atividades de preservação digital.	25
		60,4	Recursos	0	0,333	Relação custo-benefício	10	RCB01. Avaliar todos os custos ao longo do ciclo de vida dos arquivos, desde a captura até a preservação a longo prazo.	0
				90,625	0,667	Recursos Humanos	10	RH01. Definição clara de papeis e responsabilidades relacionadas às ações de preservação digital.	100
							6	RH02. Desenvolvimento de programas de treinamento contínuo para manter a equipe atualizada sobre as melhores práticas de preservação digital.	75

Fonte: elaborado pela autora

Na análise, foi possível observar que alguns requisitos fundamentais não foram atendidos. Por exemplo, o requisito *RCB01 – Avaliar todos os custos ao longo do ciclo de vida dos arquivos, desde a captura até a preservação a longo prazo*, de peso

10, teve um cumprimento de 0%, o que indica a ausência de uma análise de custos documentada ou implementada ao longo do ciclo de vida dos arquivos. O requisito *OS03 – Implementação de sistemas que sejam flexíveis para se adaptarem às mudanças tecnológicas*, de peso 8, também teve cumprimento de 0%, uma vez que o plano atual sugere apenas, de forma geral, a eventual implementação de sistemas adaptáveis, sem apresentar uma estratégia já implementada.

O requisito *OG04 – Documentação de procedimentos operacionais padrão para todas as atividades de preservação digital*, de peso 10, alcançou um cumprimento parcial de 25%. Alguns procedimentos foram documentados, especificamente as etapas de *ações de preservação, monitoramento* e *revisão*. Com base na análise dos requisitos e no percentual de cumprimento de cada um, o modelo de preservação digital alcançou uma pontuação de qualidade de 72%. No geral, essa pontuação reflete uma boa implementação do plano em diversos aspectos, indicando que a maioria dos critérios de preservação digital está sendo atendida de forma satisfatória. A pontuação é um bom indicativo de que o plano possui uma boa base, mas precisa de ajustes para atingir um nível ideal de conformidade e eficácia.

7 Objetivos concretizados

Este capítulo resume os objetivos traçados no início da pesquisa e avalia até que ponto foram atingidos durante o desenvolvimento do modelo de preservação digital proposto. A pesquisa pretendia abordar a problemática da preservação de dados digitais na área da saúde, considerando tanto o cenário tecnológico quanto o contexto regulatório e organizacional brasileiro.

a) Analisar as etapas, padrões e normas para a implantação de um plano de preservação digital

O primeiro objetivo foi cumprido ao longo dos capítulos que exploraram as normas e boas práticas reconhecidas internacionalmente, como as tratadas no Modelo de Referência OAIS e nos ciclos de vida da curadoria digital. Foram consideradas as etapas desde a criação até a preservação e descarte dos dados, de forma a assegurar a continuidade e integridade das informações médicas ao longo do tempo. A pesquisa também investigou regulamentações como a LGPD, Lei nº 13.787/2018 e resoluções do CFM, uma vez que as atividades propostas devem também estar em conformidade com as exigências legais de proteção de dados.

Além disso, foram explorados na pesquisa padrões de interoperabilidade e de metadados que são fundamentais para garantir que os dados possam ser trocados, acessados e compreendidos por diferentes sistemas e plataformas ao longo do tempo. Esse ponto também destaca a importância na adoção de padrões abertos.

Ao longo da pesquisa, também foi discutida a necessidade de um sistema de governança de dados que inclua políticas de controle de acesso, segurança da informação e responsabilidade pela gestão de dados. Isso inclui a adoção de medidas para garantir a integridade e a confidencialidade das informações, considerando a prevenção contra perdas, alterações não autorizadas e uso indevido dos dados. As estratégias e sistemas de preservação digital foram abordados com ênfase em técnicas como migração, emulação, encapsulamento, refrescamento e preservação de tecnologia.

b) Explorar o panorama atual de diretrizes de preservação digital, a partir de uma revisão da literatura

A partir da revisão da literatura realizada, que cobriu tanto o contexto nacional quanto internacional da preservação digital, o trabalho conseguiu mapear e documentar as principais diretrizes, incluindo iniciativas e práticas adotadas em diferentes, e buscou entender as lacunas da preservação digital no Brasil, considerando também o contexto da saúde.

Podemos destacar as práticas realizadas pelo IBICT, pela Rede Cariniana, pelo CONARQ e Arquivo Nacional no Brasil. Foram também exploradas iniciativas globais que contribuem para a formação de conhecimentos na área de preservação digital, como o SINPRED e o Projeto InterPARES, com foco em estratégias de preservação e cooperação entre organizações de diferentes países e que contam com relevante atuação do Brasil. O estudo também incluiu a análise de modelos de preservação como o OAIS e o Modelo Hipátia, e de redes de colaboração internacionais, como o NDSA.

No contexto brasileiro, observou-se uma escassez de políticas de preservação digital específicas para o setor público, o que representa uma oportunidade para futuras iniciativas de desenvolvimento de políticas alinhadas com o setor de saúde.

c) Identificar e caracterizar as necessidades específicas de preservação digital no contexto das bases de dados de saúde, considerando o contexto brasileiro

A pesquisa identificou as especificidades dos registros médicos eletrônicos e as necessidades relacionadas à sua preservação. Isso incluiu a análise de formatos de dados, requisitos de interoperabilidade, e os desafios de garantir a integridade e acessibilidade dos prontuários ao longo de décadas. A utilização de uma unidade fictícia, a Unidade de Saúde São Lucas, serviu para ilustrar de forma mais prática esses aspectos, apresentando um caso simulado de elaboração de um modelo de preservação digital.

Foram explorados trabalhos relevantes na área da saúde, como os publicados pela RECIIS, que forneceram o panorama de pesquisas e contribuições científicas a curadoria e preservação digital em bases de dados de saúde. Essas

referências enriqueceram a caracterização das necessidades de preservação em conjunto com o cenário de saúde digital atual no Brasil, abordando o valor do registro médico como uma fonte importante para o histórico de saúde dos pacientes e para a elaboração de políticas públicas.

d) Propor um modelo de preservação digital adaptado às características das bases de dados de saúde

Com base nos dados coletados, padrões analisados e necessidades específicas identificadas, foi desenvolvido um modelo de preservação digital direcionado às bases de dados de saúde. Este modelo contempla desde a ingestão e classificação dos dados, até as estratégias de monitoramento e mitigação de riscos, passando por recomendações para migração de formatos e adoção de tecnologias emergentes. Além disso, o modelo estabelece práticas contínuas de avaliação e revisão, garantindo que as políticas de preservação permaneçam adequadas diante das mudanças tecnológicas e regulatórias.

Todos os objetivos específicos traçados inicialmente foram atingidos, resultando em um modelo adaptável que visa contribuir para a preservação de dados críticos na saúde brasileira. O desenvolvimento do modelo buscou integrar as melhores práticas e padrões, com ajustes ao cenário local, e priorizando a segurança, integridade e acessibilidade dos registros médicos ao longo do tempo.

8 Limitações e trabalhos futuros

Nesta seção, discutimos as limitações do presente estudo e propomos direções para trabalhos futuros que possam contribuir para a evolução do modelo de preservação digital em bases de dados na área da saúde. A principal limitação desse estudo está relacionada à implementação do modelo em um ambiente real. Devido às restrições de tempo e recursos, o modelo proposto foi desenvolvido e avaliado principalmente em termos conceituais, sem uma experimentação prática em larga escala. Isso significa que não foi possível validar totalmente a efetividade do modelo em situações reais de preservação de dados de saúde, o que representa uma limitação significativa no que diz respeito à comprovação de sua viabilidade prática.

Além disso, a pesquisa não abordou aspectos econômicos detalhados para a implementação do modelo em diferentes instituições de saúde, o que pode representar um fator limitante para sua adoção. A análise de custo-benefício é essencial para avaliar a viabilidade do modelo de preservação digital, considerando as limitações orçamentárias das instituições públicas de saúde. Trabalhos futuros podem incluir a implementação e validação prática do modelo em ambientes reais, com diferentes perfis de instituições de saúde, a fim de avaliar sua efetividade, viabilidade e limitações em situações concretas. A aplicação do modelo em ambientes distintos também permitirá a adaptação às especificidades de diferentes sistemas de saúde, incluindo variações regionais e institucionais.

Uma outra direção importante para pesquisas futuras é a investigação de aspectos econômicos e a realização de uma análise de custo-benefício detalhada, considerando os custos de infraestrutura, treinamento e manutenção de sistemas necessários para a implementação do modelo. Esses estudos poderiam oferecer uma visão mais clara sobre a viabilidade financeira do modelo e facilitar sua adoção em larga escala. Além disso, é pertinente explorar novas tecnologias emergentes no campo da preservação digital, como *blockchain* (XAVIER; GOTTSCHALG-DUQUE, 2021; RABELO, 2023; BARBOSA; PEROZINI; HERMEIRO, 2023; FLORES, 2024) e soluções baseadas em inteligência artificial (HARISANTY et al., 2024), que possam contribuir para o aprimoramento das técnicas de preservação e na garantia de maior segurança e automação dos processos.

9 Conclusão

Ao longo desta pesquisa, foi proposto um modelo de preservação digital para a área da saúde, visando garantir a integridade, acessibilidade e segurança dos dados médicos ao longo do tempo. Este trabalho destaca a importância de uma abordagem adaptável que responda aos desafios específicos do setor de saúde, como as regulamentações de sigilo, a rápida obsolescência tecnológica e a diversidade de dados clínicos e administrativos que compõem os registros médicos.

O modelo apresentado é, antes de tudo, uma proposta e representa uma contribuição para com as pesquisas e práticas de preservação digital no contexto da saúde no Brasil. Ele foi desenvolvido considerando as regulamentações locais e as demandas das instituições de saúde, mas permanece flexível e pode – e deve – ser adaptado a diferentes contextos e necessidades institucionais. Com um conjunto de etapas estruturadas, o modelo pretende oferecer uma estratégia para assegurar que os dados médicos permaneçam acessíveis e íntegros, independentemente das evoluções tecnológicas.

Espera-se que esta proposta sirva de base para novas pesquisas e inspire a criação de modelos específicos que atendam às diversas realidades institucionais e regionais do Brasil. Assim, ao mesmo tempo em que proporciona uma ferramenta prática para as instituições de saúde, esta pesquisa desenvolve conhecimento e prática em preservação digital, que poderá ser expandida e aperfeiçoada conforme o setor de saúde continua a se digitalizar e evoluir.

Referências

ACADEMIA BRASILEIRA DE LETRAS. **Nato-digital**. Disponível em: <https://www.academia.org.br/nossa-lingua/nova-palavra/nato-digital>. Acesso em: 08 out. 2024.

AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR. **TISS - Padrão para Troca de Informação de Saúde Suplementar**. Disponível em: <https://www.gov.br/ans/pt-br/assuntos/prestadores/padrao-para-troca-de-informacao-de-saude-suplementar-2013-tiss>. Acesso em: 18 maio. 2024.

ALMEIDA, R. O. de; SILVA, R. C. L. da; COSTA, M. V. da S. de B. Preservação da memória na área de ciências da saúde. **Revista Brasileira de Biblioteconomia e Documentação**, v. 13, p. 2657–2672, 2017.

ARAÚJO, D. G. de; LLARENA, M. A. A.; SIEBRA, S. de A.; DIAS, G. A. Contribuições para a gestão de dados científicos: análise comparativa entre modelos de ciclo de vida dos dados. **Liinc em Revista**, v. 15, n. 2, 2019. Disponível em: <https://revista.ibict.br/liinc/article/view/4686>. Acesso em: 26 abr. 2024.

ARQUIVO NACIONAL. **Arquivo Nacional disponibiliza tutoriais para o uso das ferramentas AtoM e Archivematica**. Disponível em: <https://www.gov.br/arquivonacional/pt-br/canais_atendimento/imprensa/copy_of_noticias/arquivo-nacional-disponibiliza-tutoriais-para-o-uso-das-ferramentas-atom-e-archivematica>. Acesso em: 6 nov. 2024.

ARQUIVO PÚBLICO DO ESTADO DE SÃO PAULO. **Política de gestão e preservação de documentos digitais**. São Paulo: Arquivo Público do Estado de São Paulo, 2022. p. 129.

ARTEFACTUAL SYSTEMS INC. **Archivematica documentation | Archivematica: open-source digital preservation system**. Disponível em: <https://www.archivematica.org/pt-br/docs/archivematica-1.16/>. Acesso em: 6 nov. 2024.

ASSOCIATION OF RESEARCH LIBRARIES; CENTER FOR MEDIA & SOCIAL IMPACT; PROGRAM ON INFORMATION JUSTICE AND INTELLECTUAL PROPERTY. **Code of Best Practices in Fair Use for Software Preservation**. 2018. Disponível em: <https://www.arl.org/wp-content/uploads/2019/03/Code-of-Best-Practices-in-Fair-Use-for-Software-Preservation.pdf>. Acesso em: 23 maio 2024.

BAGGIO, C. C.; FLORES, D. Estratégias, critérios e políticas para preservação de documentos digitais em arquivos. **Ciência da Informação**, v. 41, n. 2/3, p. 58–71, 2015.

BALL, A. **Review of Data Management Lifecycle Models**. University of Bath, UK United Kingdom, 2012. Disponível em: <https://doi.org/doi:10.5060/D2251G48>. Acesso em: 26 abr. 2024.

BARATEIRO, J. et al. Designing Digital Preservation Solutions: A Risk Management-Based Approach. **International Journal of Digital Curation**, v. 5, n. 1, p. 4–17, 22 jun. 2010.

BARBOSA, T. S. G; PEROZINI, S.; FLORES, D. Uma reflexão sobre possibilidades do uso do blockchain na Arquivologia. **Revista do Arquivo Público do Estado do Espírito Santo**, [S. l.], v. 7, n. 13, 2024. Disponível em: https://periodicos.ufes.br/revapees/article/view/43657. Acesso em: 19 nov. 2024.

BARBOSA, W. L.; SHAYER LYRA, R. Governança de dados. **Escola Nacional de Administração Pública (ENAP)**, 2021.

BECKER, C. et al. Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. **International Journal on Digital Libraries**, v. 10, n. 4, p. 133–157, dez. 2009.

BOTÃO, A. V. R. **Metadados para tratamento de imagens médicas como objeto de ensino e aprendizagem com fins de reuso**. Tese (Doutorado em Ciência da Informação) – Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia. Rio de Janeiro, 117 p. 2019.

BOTE, J.; FEIJOO, B. F.; RUIZ, S. Digital preservation cost: a cost accounting approach. **The Learning Organization**, v. 20, n. 6, p. 419–432, 2013.

BRASIL. Constituição da República Federativa do Brasil. Brasília: Senado, 1988.

BRASIL. **Estratégia de Saúde Digital para o Brasil**. Brasília: Ministério da Saúde (MS), Departamento de Informática do SUS, 2020a. Disponível em: <http://bvsms.saude.gov.br/bvs/publicacoes/estrategia_saude_digital_Brasil.pdf>. Acesso em: 29. abr. 2024.

BRASIL. Ministério da Saúde. Comitê Gestor da Estratégia e-Saúde. **Estratégia e-Saúde para o Brasil**. Brasília. 2017. Disponível em: <https://www.conasems.org.br/wp-content/uploads/2019/02/Estrategia-e-saude-para-o-Brasil.pdf>. Acesso em: 03 maio. 2024.

BRASIL. Ministério da Saúde. Secretaria de Atenção Primária à Saúde. e-SUS Atenção Primária à Saúde: Manual do Sistema com Prontuário Eletrônico do Cidadão PEC. Brasília. 2023.

BRASIL. **Lei nº 8.080, de 19 de Setembro de 1990**. Dispõe sobre as condições para a promoção, proteção e recuperação da saúde, a organização e o funcionamento dos serviços correspondentes e dá outras providências. Brasília: Senado. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/l8080>. Acesso em: 15 out. 2024.

BRASIL. **Lei nº 8.159, de 8 de Janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Brasília: Senado. Disponível em: <https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 26 abr. 2024.

BRASIL. **Lei nº 10.406, de 10 de Janeiro de 2002**. Institui o Código Civil. Brasília, Senado. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm>. Acesso em: 26 abr. 2024.

BRASIL. **Lei nº 12.527, de 18 de Novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º [...]. Brasília, Senado, 2011a. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em: 26 abr. 2024.

BRASIL. **Lei nº 12.550, de 15 de Dezembro de 2011**. Autoriza o Poder Executivo a criar a empresa pública denominada Empresa Brasileira de Serviços Hospitalares – EBSERH [...]. Brasília, Senado, 2011b. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12550.htm>. Acesso em: 05 jun. 2024.

BRASIL. **Lei nº 13.709, de 14 de Agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, Senado, 2018a. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em: 26 abr. 2024.

BRASIL. **Lei nº 13.787, de 27 de Dezembro de 2018**. Dispõe sobre a digitalização e a utilização de sistemas informatizados para a guarda, o armazenamento e o manuseio de prontuário de paciente. Brasília, Senado, 2018b. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13787.htm>. Acesso em: 26 abr. 2024.

BRASIL. **Lei nº 14.063, de 23 de Setembro de 2020**. Dispõe sobre o uso de assinaturas eletrônicas em interações com entes públicos, em atos de pessoas jurídicas e em questões de saúde e sobre as licenças de softwares desenvolvidos por entes públicos [...]. Brasília, 2020. Disponível em: < https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/l14063.htm>. Acesso em: 30 out. 2024.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS. **Plano de ação, monitoramento e avaliação da estratégia de Saúde Digital para o Brasil 2019-2023**. Brasília, DF: MS, 2020b. Disponível em: <https://saudedigital.saude.gov.br/wp-content/uploads/2020/04/PAMA-Saude-digital.pdf>. Acesso em: 03 maio. 2024.

BRASIL. Ministério da Saúde. **Sobre o Datasus**. Disponível em: <https://datasus.saude.gov.br/sobre-o-datasus>. Brasília, DF. 2023ª. Acesso em: 03 maio. de 2023.

BRAYNER, A. A. Curadoria digital: novos modelos de participação pública na descrição de conteúdos em instituições culturais. **Revista Ibero-Americana de Ciência da Informação**, [S.l.], v. 12, n. 1, p. 53–65, 2018. DOI: 10.26512/rici.v12.n1.2019.10521. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/10521>. Acesso em: 8 out. 2023.

BREEDING, M. Ongoing Challenges in Digitization. **Computers in Libraries**, v. 34, n. 09, p. 16–18, 2014.

BROWN, A. Practical digital preservation: a how-to guide for organizations of any size. London: Facet Pub, 2013.

BUARQUE, M. D.; MACHADO, J. G. N.; PONTES, E. B. Plano de Preservação Digital da VideoSaúde: estratégias para a gestão de documentos audiovisuais. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 14, n. 3, 29 set. 2020.

CANADIAN HERITAGE INFORMATION NETWORK (CHIN). **Digital Preservation Inventory Template for Museums**. Disponível em: <https://www.canada.ca/en/heritage-information-network/services/digital-preservation/inventory-template-museums.html>. Acesso em: 5 out. 2024.

CAPLAN, P. **Entendendo o PREMIS**. Washington, D.C.: The Library of Congress, 2009. Disponível em: <https://loc.gov/standards/premis/understandingPREMIS>. Acesso em: 27 out. 2024.

CASTRO, F. F.; SANTOS, P. L. V. A. C. Metadados em ciência da informação: trajetória científica no Brasil. **XIX Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação**, v. 24, n. 2, 19 abr. 2018.

CAVALCANTI MOREIRA, F. **Proposta de modelo de preservação digital para repositórios digitais**. 94 f. Dissertação (Mestrado em Gestão de Unidades de Informação) - Universidade do Estado de Santa Catarina, Florianópolis, 2017. Disponível em: <https://www.udesc.br/arquivos/faed/id_cpmenu/1440/dissertacao_fernando_15689008448301_1440.pdf>. Acesso em: 26 abr. 2024.

CCSDS. **Reference Model for an Open Archival Information System (OAIS).** Consultative Committee for Space Data Systems (CCSDS), 2012.

CHAPMAN, S. Chapter 2: Managing Digitization. **Library Technology Reports**, v. 40, n. 5, p. 13–21, 2004.

CONARQ. Conselho Nacional de Arquivos. **Carta para a Preservação do Patrimônio Arquivístico Digital**. Brasília: Conselho Nacional de Arquivos, 2004.

CONSELHO FEDERAL DE MEDICINA. **Código de Ética Médica**: Resolução CFM nº 2.217, de 27 de setembro de 2018, modificada pelas Resoluções nº 2.222/2018 e 2.226/2019. Brasília: Conselho Federal de Medicina, 2019.

CONSELHO FEDERAL DE MEDICINA. **Resolução CFM Nº 1.821/2007**, de 23 de novembro de 2007, modificada pela Resolução CFM nº 2.218/2018. Brasília, DF: Conselho Federal de Medicina, 2007.

CONSTANTOPOULOS, P.; DALLAS, C. Aspects of a digital curation agenda for cultural heritage. In: **IEEE International Conference on Distributed Human-Machine Systems**, 2007.

CONSTANTOPOULOS, P.; DALLAS, C. et al. DCC&U: An Extended Digital Curation Lifecycle Model. In: **The International Journal of Digital Curation**, 2009. Disponível em: <https://doi.org/10.2218/ijdc.v4i1.76>. Acesso em: 21 mar. 2024.

COPTR. 2021. Disponível em: <https://coptr.digipres.org/index.php/Main_Page>.

D'AGOSTINO, M. et al. Estrategia para la gobernanza de datos abiertos de salud: un cambio de paradigma en los sistemas de información. **Revista Panamericana de Salud Pública**, v. 41, p. 1, 23 mar. 2017.

D'AGOSTINO, M. et al. Prontuários médicos do paciente: a digitalização não é mais uma opção e deve ser uma obrigação. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 14, n. 3, 29 set. 2020.

DAMA International. **DAMA-DMBOK: Data Management Body of Knowledge: 2nd Edition**, Revised. Technics Publications, 2ª edição, 4 de julho de 2017.

DELANEY, B.; JONG, A. DE. Media Archives and Digital Preservation: Overcoming Cultural Barriers. **New Review of Information Networking**, v. 20, n. 1-2, p. 73–89, 3 jul. 2015.

DIGITAL CURATION CENTRE. **History of the DCC**. Disponível em: https://www.dcc.ac.uk/about/history-dcc. Acesso em: 8 out. 2023.

DIGITAL PRESERVATION COALITION. **Digital curation: digital archives, libraries and e-science**. Londres, 19 out. 2001. Disponível em: www.dpconline.org/events/past-events/digital-curation. Acesso em: 8 out. 2023.

DIGITAL PRESERVATION COALITION. **Digital Preservation Handbook**. 2. ed. Digital Preservation Coalition, 2015. Disponível em: <http://handbook.dpconline.org/>. Acesso em: 30 set. 2024.

DIGITAL PRESERVATION COALITION. **Digital Preservation Policy Toolkit**. 2023. Disponível em: <https://www.dpconline.org/digipres/implement-digipres/policy-toolkit>. Acesso em: 8 out. 2023.

DIGITAL PRESERVATION COALITION. Digital Preservation Coalition Rapid Assessment Model (DPC RAM). 22 mar. 2024.

DIGITALNZ. **Make it Digital**. 2006a. Disponível em: <https://digitalnz.org/make-it-digital/getting-started-with-digitisation>. Acesso em: 8 out. 2023.

DIGITALNZ. **Our Strategy**. 2006b. Disponível em: <https://digitalnz.org/about/our-strategy>. Acesso em: Acesso em: 8 out. 2023.

DATA PROTECTION OFFICER (DPO). **European Data Protection Supervisor**. Disponível em: <https://www.edps.europa.eu/data-protection/data-protection/reference-library/data-protection-officer-dpo_en>.

ELIAS, P. R. **A conservação de fitas magnéticas**. Disponível em: <https://webinsider.com.br/2015/06/29/a-conservacao-de-fitas-magneticas/>. Acesso em: 13 out. 2024.

ESCUDEIRO, P.; BIDARRA, J. Quantitative Evaluation Framework (QEF). **Revista Iberica de Sistemas y Tecnologias de Informacion**, n.1, p. 16–27, 2008.

Estratégia Saúde da Família. Disponível em: <https://www.gov.br/secom/pt-br/acesso-a-informacao/comunicabr/lista-de-acoes-e-programas/estrategia-saude-da-familia>. Acesso em: 25 ago. 2024.

EYNDEN, V. V. D**. Data Life Cycle & Data Management Planning**. Colchester: Research Data Management Team, 2013. Disponível em: <https://dam.ukdataservice.ac.uk/media/187718/dmplanningdm24apr2013.pdf>.

FARIA, M. M. DE; SILVA, T. C. Estudos de custos para preservação digital e repositório digital confiável. **Revista Brasileira de Preservação Digital**, v. 5, n. e024002, 28 jun. 2024.

FERNANDES, M. Hospital São João em parceria com a KEEP SOLUTIONS lidera projeto pioneiro na preservação de informação clínica - KEEP SOLUTIONS. Disponível em: <https://www.keep.pt/2021/06/15/hospital-sao-joao-em-parceria-

com-a-keep-solutions-lidera-projeto-pioneiro-na-preservacao-de-informacao-clinica/>. Acesso em: 4 nov. 2024.

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

FERREIRA, M.; SARAIVA, R.; RODRIGUES, E. **Estado da Arte em Preservação Digital**. 2012. Disponível em: <http://hdl.handle.net/1822/17049>. Acesso em: 3 abr. 2024.

FHIR for FAIR - FHIR Implementation Guide: **Metadata and data**. 2020. Disponível em: <https://build.fhir.org/ig/HL7/fhir-for-fair/metadata.html>. Acesso em: 3 abr. 2024

FORMENTON, D.; GRACIOSO, L. DE S. Preservação digital: desafios, requisitos, estratégias e produção científica. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 18, n. e020012, 8 jun. 2020.

FORMENTON, D.; GRACIOSO, L. DE S.; CASTRO, F. F. DE. Revisitando a preservação digital na perspectiva da ciência da informação. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 13, n. 1, p. 170–191, 30 jan. 2015.

FRASÃO, G.; RIBEIRO, K. **Atenção Primária e Atenção Especializada: Conheça os níveis de assistência do maior sistema público de saúde do mundo**. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/noticias/2022/marco/atencao-

primaria-e-atencao-especializada-conheca-os-niveis-de-assistencia-do-maior-sistema-publico-de-saude-do-mundo>. Acesso em: 25 ago. 2024.

GRÁCIO, J. C. A. Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior. São Paulo: Cultura Acadêmica, 2012.

GRAELLS, M. T. A importância da preservação digital para os sistemas de saúde. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 14, n. 3, 2020.

HARISANTY, D. et al. Cultural heritage preservation in the digital age, harnessing artificial intelligence for the future: a bibliometric analysis. **Digital Library Perspectives**, v. 40, n. 4, p. 609–630, 2024.

HARVEY, D. R.; OLIVER, G. **Digital curation**. Chicago: Ala Neal-Schuman, An Imprint of The American Library Association, 2016.

HEARD, S. **openEHR – What is openEHR**. Disponível em: <https://openehr.org/about/what_is_openehr>. Acesso em: 9 set. 2024.

HEALTHCARE INFORMATION AND MANAGEMENT SYSTEMS SOCIETY. **Dictionary of healthcare information, technology terms, acronyms and organizations**. Chicago: Healthcare Information and Management Systems Society, 2006.

HEDSTROM, M. Digital Preservation: A Time Bomb for Digital Libraries. **Computers and the Humanities**, v. 31, n. 3, p. 189–202, 1997.

HERMEIRO, A. C. C. **A cadeia de custódia da prova digital: O uso da Tecnologia Blockchain como forma de preservação**. Dissertação (Mestrado em Ciências Jurídico-Forenses) – Universidade de Coimbra, Faculdade de Direito. Coimbra, 64 p. 2023.

HIGGINS, S. The DCC Curation Lifecycle Model. In: **The International Journal of Digital Curation**, 2008. Disponível em: <https://doi.org/10.2218/ijdc.v3i1.48>. Acesso em: 8 out. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 14721:2012 Space data and information transfer systems – Open archival information system (OAIS) – Reference model. Geneva, 2012.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/TR 20514:2005 Health informatics — Electronic health record — Definition, scope and context. Genebra, 2005.

INTERPARES. **Projeto InterPARES - International Research on Permanent Authentic Records in Electronic Systems**. Disponível em: <http://interparestrust.org/>. Acesso em: 8 out. 2023.

INTERPARES PROJECT. **Diretrizes do produtor: a elaboração e a manutenção de materiais digitais**. Tradução: Arquivo Nacional; Tradução: Câmara dos Deputados. Brasília, DF: Câmara dos Deputados, 2011.

Introdução | RNDS. Disponível em: <https://rnds-guia.saude.gov.br/docs/introducao>. Acesso em: 3 maio. 2024.

JAPIASSU, H.; MARCONDES, D. **Dicionário básico de filosofia**. Rio De Janeiro: J. Zahar, 2001.

JÚNIOR, L. P. DA S.; MOTA, V. G. DA. Políticas de preservação digital no Brasil: características e implementações. **Ciência da Informação**, v. 41, n. 1, 2014.

KENNEY, A. R. *et al*. Digital Preservation Management: Digital Preservation Management Workshops and Tutorial. Disponível em: <https://dpworkshop.org/index.php/dpm-eng.html>.

KHURANA, A.; ROSENTHAL, S. R. Towards Holistic "Front Ends" In New Product Development. **Journal of Product Innovation Management**, v. 15, n. 1, p. 57–74, jan. 1998.

KIATAKE, G. G. L. et al. **Manual de Certificação de Sistemas de Registro Eletrônico em Saúde**. Resolução CFM nº 1821/2007. 2020. Disponível em: <https://www.sbis.org.br/certificacao/Manual_Certificacao_S-RES_SBIS_v5-0.pdf>. Acesso em: 9 set. 2024.

KOEN, P. et al. Providing Clarity and A Common Language to the "Fuzzy Front End". **Research-Technology Management**, v. 44, n. 2, p. 46–55, mar. 2001.

LAVOIE, B. **The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition)**. Digital Preservation Coalition, 2014. Disponível em: <http://dx.doi.org/10.7207/twr14-02>. Acesso em: 12 dez. 2023.

LEE, K.-H. et al. The State of the Art and Practice in Digital Preservation. **Journal of Research of the National Institute of Standards and Technology**, v. 107, n. 1, p. 93–106, 2002.

LEVELS OF PRESERVATION REVISIONS WORKING GROUP. **Levels of Digital Preservation Matrix V2.0**. 13 nov. 2023. Disponível em: https://osf.io/2mkwx/.

LIN, A. et al. Governança dos dados na saúde: Uso e compartilhamento. 2023.

LIRA, J.; SIEBRA, S. A. Preservação Digital: revisitando o essencial. In: SIEBRA, S. de A.; BORBA, V. da R. (Orgs.). **Preservação Digital e suas facetas**. São Carlos: Pedro & João Editores, 2021. p. 31-83.

LOCKSS. **How LOCKSS Works | LOCKSS Program**. Disponível em: <https://www.lockss.org/use-lockss/how-lockss-works>. Acesso em: 24 maio. 2024.

LUCASSEN, G. et al. The Use and Effectiveness of User Stories in Practice. **Requirements Engineering: Foundation for Software Quality**, p. 205–222, 2016.

MACHADO, M. L.; TAVARES, S. **Programa TechSUS: Governança e interoperabilidade de dados para a Saúde**. Brasília: Instituto de Estudos para Políticas de Saúde, 2023. Disponível em: <https://static.poder360.com.br/2023/03/panorama-ieps-4-techSUS-saude-digital.pdf>.

MÁRDERO ARELLANO, M. Á. Cariniana: uma rede nacional de preservação digital. **Ciência da Informação**, v. 41, n. 1, 8 abr. 2014.

MÁRDERO ARELLANO, M. Á. **Critérios para a preservação digital da informação científica**. Tese (Doutorado em Ciência da Informação). Universidade Federal de Brasília, Departamento de Ciência da Informação. Distrito Federal, 356 p. 2008.

MARQUES, P. E. et al. **Manual de Certificação para Sistemas de Registro Eletrônico em Saúde**. 2016. Disponível em: <https://www.sbis.org.br/certificacao/Manual_Certificacao_SBIS-CFM_2016_v4-2.pdf>.

MCMEEKIN, S.; CURRIE, A. DPC Digital Preservation Competency Framework. 2022.

MEIRELLES, R. F.; CUNHA, F. J. A. P. Autenticidade e preservação de Registros Eletrônicos em Saúde: proposta de modelagem da cadeia de custódia das informações orgânicas do Sistema Único de Saúde. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 14, n. 3, 2020.

MELLO, A. P. P.; MESQUITA, H.; VIEIRA, C. E. Introdução à Interoperabilidade (ePING). **Escola Nacional de Administração Pública (ENAP)**, 2015.

MELLO, J. **Proposta de modelo para a preservação e curadoria digital de objetos digitais de centros de pesquisas oncológicas**. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Programa de Pós-Graduação em Ciência da Informação. Florianópolis, 360 p. 2020.

MELLO, J.; VIANNA, W. B. Preservação digital da informação em saúde: panorama quali-quantitativo da produção científica internacional. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, [S. l.], v. 14, n. 2, 2019. Disponível em: <https://pbcib.com/index.php/pbcib/article/view/45380>. Acesso em: 11 nov. 2024.

MINISTÉRIO DA SAÚDE. Plano para implantação do padrão ISBT 128 nos serviços de hemoterapia. Brasília: Ministério da Saúde. 2012.

MINISTÉRIO DA SAÚDE. **Política Nacional de Informação e Informática em Saúde**. Brasília: Ministério da Saúde. 2016.

MINISTÉRIO DA SAÚDE. **PORTARIA GM/MS Nº 1.768.** 2021. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-gm/ms-n-1.768-de-30-de-julho-de-2021-335472332>.

MINISTÉRIO DA SAÚDE. **PORTARIA Nº 4.279, de 30 de dezembro de 2010 – Anexo**. Estabelece diretrizes para a organização da Rede de Atenção à Saúde no âmbito do Sistema Único de Saúde (SUS). 2010. Disponível em: < https://bvsms.saude.gov.br/bvs/saudelegis/gm/2010/prt4279_30_12_2010.html>.

NASCIMENTO, A. F. G. DO; ARAÚJO, L. D. DE; MÁRDERO ARELLANO, M. Á. Crise e oportunidades para a preservação digital da informação em saúde. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 14, n. 3, 29 set. 2020.

NATIONAL DIGITAL STEWARDSHIP ALLIANCE. **LOP Implementation Guide and Working Definitions**. 6 jun. 2022. Disponível em: < https://doi.org/10.17605/OSF.IO/NT8U9>. Acesso em: 18 maio. 2024.

OFFICE OF THE PRIVACY COMMISSIONER FOR PERSONAL DATA. **Guidance on Personal Data Erasure and Anonymisation**. Hong Kong: PCPD, 2011. Disponível em: <https://www.pcpd.org.hk/english/publications/files/erasure_e.pdf>. Acesso em: 31 out. 2024.

OSTERWALDER, A. **The business model ontology a proposition in a design science approach**. Tese (Doutorado em Informática Empresarial) — Universidade de Lausanne, Ecole des Hautes Etudes Commerciales. Lausanne, 172 p. 2004.

OSTERWALDER, A. et al. Value Proposition Design: How to Create Products and Services Customers Want. Somerset: Wiley, 2014.

OSTERWALDER, A.; PIGNEUR, Y.; CLARK, T. Modeling value propositions in e-Business. **ACM International Conference Proceeding Series**, 2023. 50. 429-436. 10.1145/948005.948061.

OSTERWALDER, A.; PIGNEUR, Y.; CLARK, T. **Business Model Generation: a Handbook for visionaries, Game changers, and Challengers**. Hoboken, New Jersey: Wiley, 2010.

OWENS, T. **The theory and craft of digital preservation**. Baltimore, Maryland: Johns Hopkins University Press, 2018.

PADRÃO, M. **Qual a validade de um HD? Cuidado para ele não "vencer" e sumir com dados**. Disponível em: <https://www.uol.com.br/tilt/noticias/redacao/2020/10/29/qual-o-prazo-de-validade-de-um-hd-cuidado-ele-pode-expirar.htm>. Acesso em: 13 out. 2024.

PERSONAL DATA PROTECTION COMMISSION SINGAPORE. **Guide to Basic Anonymisation**. Hong Kong: PDPC, 2022. Disponível em: <https://www.pdpc.gov.sg/>. Acesso em: 19 set. 2024.

PIETROBON, L.; PRADO, M. L. DO; CAETANO, J. C. Saúde suplementar no Brasil: o papel da Agência Nacional de Saúde Suplementar na regulação do setor. **Physis: Revista de Saúde Coletiva**, v. 18, n. 4, p. 767–783, 2008.

PINTO, D.; CARNEIRO, R. Análise das Políticas de Preservação Digital para Documentos de Arquivo das Instituições de Ensino Superior do Brasil. **Brazilian Journal of Information Science research trends**, v. 18, p. e024023–e024023, 30 jul. 2024.

PINTO, V. B.; SALES, O. M. M. Proposta de aplicabilidade da preservação digital ao prontuário eletrônico do paciente. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 15, n. 2, p. 489, 7 abr. 2017.

PLALE, B.; KOUPER, I. The Centrality of Data. **Data Analytics for Intelligent Transportation Systems**, p. 91–111, 2017.

POLO ARCHIVISTICO DELL'EMILIA-ROMAGNA. **ParER - Emilia-Romagna Digital Archive Centre**. Disponível em: <https://poloarchivistico.regione.emilia-romagna.it/english>. Acesso em: 8 out. 2024.

PREMIS EDITORIAL COMMITTEE. **PREMIS Data Dictionary for Preservation Metadata**. The Library of Congress, 2015.

PRODANOV, C. C.; FREITAS, E. C. de. Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico. 2. ed. Editora Feevale, 2013.

QUEENSLAND GOVERNMENT. **File formats for long-term digital records**. Disponível em: <https://www.forgov.qld.gov.au/information-and-communication-technology/recordkeeping-and-information-management/recordkeeping/store-protect-and-care-for-records/file-formats-for-long-term-digital-records>. Acesso em: 5 nov. 2024.

RABELO, N. B. **Uso de blockchain nos arquivos: da autenticidade à autenticação de documentos**. 2023. 145 f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto de Arte e Comunicação Social, Universidade Federal Fluminense, Niterói, 2023.

REDE CARINIANA. **Rede Brasileira de Preservação Digital**. Disponível em: <https://cariniana.ibict.br/>. Acesso em: 8 out. 2023.

RICH, N.; HOLWEG, M. Value Analysis, Value Engineering. [s.l.] **Lean Enterprise Research Centre**, 2000. Disponível em: <https://ivma.org.au>. Acesso em: 7 maio. 2024.

RILEY, J. **Understanding metadata**. Baltimore: National Information Standards Organization, 2017.

ROCHA, R. P. da; PIRES, C. de O. Finalidade e Atividades da Curadoria Digital na Perspectiva de sua Implantação em uma Instituição. **Brazilian Journal of Information Science: research trends**, Marília, SP, v. 14, n. 4 - out-dez, p. e020012, 2020.

SAATY, T. L. Decision making with the analytic hierarchy process. **International Journal of Services Sciences**, v. 1, p. 83–98, 2008.

SALES, O. M. M. **Preservação digital nas Ciências da Saúde: modelo de metadados para preservação do prontuário eletrônico do paciente**. Tese (Doutorado em Ciência da Informação) – Centro de Ciências Sociais Aplicadas, Universidade Federal da Paraíba. João Pessoa, p. 247. 2022.

SALES, O. M. M.; BENTES PINTO, V. Tecnologias digitais de informação para a saúde: revisando os padrões de metadados com foco na interoperabilidade. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 13, n. 1, 29 mar. 2019.

SANTOS, H. M.; FLORES, D. As vulnerabilidades dos documentos digitais: Obsolescência tecnológica e ausência de políticas e práticas de preservação digital. **Biblios Journal of Librarianship and Information Science**, n. 59, p. 45–54, 2 jul. 2015a.

SANTOS, H. M.; FLORES, D. Estratégias de preservação digital para documentos arquivísticos: uma breve reflexão. **Cadernos BAD**, v. 1, n. 1, p. 87–101, 17 jul. 2015b.

SANTOS, H. M.; FLORES, D. Os impactos da obsolescência tecnológica frente à preservação de documentos digitais. **Brazilian Journal of Information Science: research trends**, Marília, SP, v. 11, n. 2, 2017. DOI: 10.36311/1981-1640.2017.v11n2.04.p28. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/5550.> Acesso em: 21 out. 2024.

SANTOS, H. M.; FLORES, D. Preservação de documentos arquivísticos digitais: reflexões sobre o uso de padrões abertos nos acervos. **Investigación Bibliotecológica Archivonomía Bibliotecología e Información**, v. 32, n. 74, p. 35–35, 25 jan. 2018.

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade**, v. 22, n. 3, 2012. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/12224>. Acesso em: 26 abr. 2024.

SHAHPORI, R.; DOIG, C. Systematized Nomenclature of Medicine–Clinical Terms direction and its implications on critical care. **Journal of Critical Care**, v. 25, n. 2, p. 364.e1–364.e9, jun. 2010.

SIEBRA, S. A. Curadoria Digital: uma área em expansão. **Archeion Online**, [S. l.], v. 6, n. 2, p. p.1–6, 2019. DOI: 10.22478/ufpb.2318-6186.2019v6n2.47089. Disponível em: <https://periodicos.ufpb.br/index.php/archeion/article/view/47089>. Acesso em: 26 abr. 2024.

SIEBRA, S. A.; SILVA, F. de M. O. Da Preservação Digital à Curadoria Digital. In: SIEBRA, S. de A.; BORBA, V. da R. (Orgs.). **Preservação Digital e suas facetas**. São Carlos: Pedro & João Editores, 2021. p. 265-302.

SIEBRA, S. A. et al. Projetos de curadoria digital: um relato de experiências. **Bibliotecas: Anales de Investigación**, v. 14, n. 2, p. 164–178, 2018.

SILVA, F. de M. O.; SIEBRA, S. de A. Aplicação do DCC&U para Curadoria de Objetos Culturais Digitais. **Informação & Tecnologia**, v. 4, n. 2, p. 26–45, 2018. Disponível em: <https://periodicos.ufpb.br/index.php/itec/article/view/40520>. Acesso em: 26 abr. 2024.

SILVA, F. P. **Gestão de dados em organizações: uma análise da aplicação da preservação e curadoria digital no corpo de conhecimento do DMBOK**. Tese (Mestrado em Ciência da Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul. Porto Alegre, p. 127. 2023.

SILVA, W.; FLORES, D. Política arquivística de preservação digital: um estudo sobre sua aplicabilidade em instituições públicas federais. **Perspectivas em Ciência da Informação**, v. 23, n. 3, p. 144–166, 2018.

SOCIETY OF AMERICAN ARCHIVISTS. **Digitize**. Disponível em: <https://dictionary.archivists.org/entry/digitize.html>. Acesso em: 08 out. 2024.

STRASSER, C. et al. **Primer on Data Management: What you always wanted to know**. UC Office of the President: California Digital Library, 2012. Disponível em: <https://escholarship.org/uc/item/7tf5q7n3>. Acesso em: 3 out. 2023.

SULLIVAN, S. J. An archival/records management perspective on PDF/A. **Records Management Journal**, v. 16, n. 1, p. 51–56, jan. 2006.

UNIÃO EUROPEIA. Parlamento Europeu; Conselho Europeu. **Regulamento (UE) nº 2016/679, de 27 de abril de 2016**. Regulamento Geral sobre a Proteção de Dados: relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE. Disponível em: <https://eur-lex.europa.eu/legal-content/>. Acesso em: 02 maio. 2024.

XAVIER, A. C. C.; GOTTSCHALG-DUQUE, C. A contribuição do arquivista para prontuários eletrônicos do paciente frente à tecnologia blockchain. **Ciência da Informação Express**, v. 2, n. v. 2, 2021.

WEBB, C. **Guidelines for the preservation of digital heritage**. UNESCO, 2003. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000130071>. Acesso em: 3 out. 2023.

WOODALL, T. Conceptualising "value for the customer": an attributional, structural and dispositional analysis. **Academy of Marketing Science Review**, v. 2003, n. 12, 1 jan. 2003.

WORLD HEALTH ORGANIZATION. **Global strategy on digital health 2020-2025.** [s.l.] World Health Organization, 2021.

ZABBIX. **Zabbix Solution in Healthcare & Medicine**. Disponível em: <https://www.zabbix.com/healthcare_and_medicine>. Acesso em: 8 nov. 2024.

ZEITHAML, V. A. Consumer Perceptions of price, quality, and value: a means-end Model and Synthesis of Evidence. **Journal of Marketing**, v. 52, n. 3, p. 2–22, 1988.