



FUNDAÇÃO OSWALDO CRUZ
PROGRAMA DE PRESERVAÇÃO
DIGITAL DE ACERVOS DA FIOCRUZ

Fundação Oswaldo Cruz

Programa de Preservação
Digital de Acervos da Fiocruz

Rio de Janeiro
2020

FUNDAÇÃO OSWALDO CRUZ

Presidência

Nísia Trindade Lima

Vice-Presidência de Ambiente, Atenção e Promoção da Saúde

Marco Antonio Carneiro Menezes

Vice-Presidência de Pesquisa e Coleções Biológicas

Rodrigo Correa de Oliveira

Vice-Presidência de Educação, Informação e Comunicação

Cristiani Vieira Machado

Vice-Presidência de Produção e Inovação em Saúde

Marco Aurelio Krieger

Vice-Presidência de Gestão e Desenvolvimento Institucional

Mario Santos Moreira

Elaboração:

Grupo de Trabalho responsável pela elaboração do Programa de Preservação Digital de Acervos da Fiocruz, criado pela Portaria 061/2014-PR de 16/01/2014 e Portaria 405/2017-PR de 23/03/2017

Coordenação:

Marcos José de Araújo Pinheiro (COC e Coordenador do Comitê do Preservo)

Coordenação Executiva:

Marcus Vinícius Pereira da Silva (COC) e Rodrigo Ferrari (VPEIC)

Consultoria e Desenvolvimento:

Andrea Gonçalves do Nascimento e Marco Dreer

Consultoria Externa:

Miguel Ángel Márdero Arellano (IBICT)

Relatoria:

Aline Souto (VPPCB), Luciana Danielli (ICICT), Marco Dreer (Consultoria e Desenvolvimento), Marcus Vinícius Pereira da Silva (COC), Rodrigo Ferrari (VPEIC), Vanessa Arruda (VPEIC)

Grupo de Trabalho 2014-2017:

André Luís Nogueira Campos (COGTIC)
Barbara Cristina Euzébio Pereira
Dias de Oliveira (IOC)
Cleber Belmiro (COC)
Gilmar Muniz (IOC)
Ivone Sá (COC)
Leonardo de Oliveira (ICICT)
Manuela da Silva (VPPLR)
Marcello Santana (COC)
Marcos José de Araújo Pinheiro (COC)
Mauro Mauricio Carneiro Campello (ICICT)
Rodrigo Ferrari (VPEIC)
Thiago da Fonseca Carelli (COGTIC)

Grupo de Trabalho 2017-2019:

Aline da Silva Alves (ICICT)
Álvaro Funcia Lemme (COGTIC)
Barbara Cristina Euzébio Pereira
Dias de Oliveira (IOC)
Carolina Sacramento (COC)
Caroline Perez Ghirardelli (VPPCB)
Claudete Fernandes de Queiroz (ICICT)
Cleber Belmiro (COC)
Gilmar Muniz (IOC)
Inês Santos Nogueira (COC)
Ivone Sá (COC)
João Guilherme Nogueira Machado (ICICT)
Karina Veras Praxedes dos Santos Gagliano (COC)
Laurinda Rosa Maciel (COC)
Manuela da Silva (VPPCB)
Marcello Santana (COC)
Marcos José de Araújo Pinheiro (COC)
Marise Terra Lachini (COC)
Mauro Mauricio Carneiro Campello (ICICT)
Rodrigo Ferrari (VPEIC)
Rodrigo Mexas (ICICT)
Thiago da Fonseca Carelli (COGTIC)

Grupo de Trabalho 2019-2020:

Aline Lacerda (COC)
Aline Souto (VPPCB)
Andrea Gonçalves do Nascimento (Consultoria e Desenvolvimento)
Claudete Queiroz (ICICT)
Daniel Moreira (COC)
Eliane Pontes (ICICT)
Erica Lopes (COC)
Geraldo Sorte (COGETIC)
Inês Nogueira (COC)
Jeferson Mendonça (COC)
João Guilherme Machado (ICICT)
Karina Praxedes (SIGDA/COC)
Luciana Danielli (ICICT)
Manuela da Silva (VPPCB)
Marcelo Santana (COC)
Marco Dreer (Consultoria e Desenvolvimento)
Marcos José de Araújo Pinheiro (COC)
Marcus Vinícius Pereira da Silva (COC)
Misael Araújo (COGETIC)
Regina Simões (COC)
Rodrigo Ferrari (VPEIC)
Rodrigo Mexas (ICICT)
Tania Santos (ICICT)
Thiago Carelli (COGETIC)
Vanessa Arruda (VPEIC)
Viviane Veiga (ICICT)

Capa, projeto gráfico e diagramação:

Silmara Mansur



Atribuição não comercial (CC BY-NC): Essa licença permite que outros remixem, adaptem e criem a partir deste material para fins não comerciais, e, embora materiais derivados tenham de atribuir o devido crédito e não possam ser usados para fins comerciais, os usuários não são obrigados a licenciar as materiais derivados sob os mesmos termos.

Licença disponível em:

https://creativecommons.org/licenses/by-nc/4.0/deed.pt_BR

Obras de terceiros e conceitos utilizados neste material são de inteira responsabilidade da área técnica.

O conteúdo desta e de outras obras da Fiocruz pode ser acessado na página: www.arca.fiocruz.br

Elaboração, distribuição e informações:

FUNDAÇÃO OSWALDO CRUZ

Av. Brasil, 4365, Manguinhos, CEP 21040-900 – Rio de Janeiro/RJ

Site: www.fiocruz.br

Impresso no Brasil / Printed in Brazil

Ficha Catalográfica

F981 Fundação Oswaldo Cruz.

Programa de preservação digital de acervos da Fiocruz / Fundação Oswaldo Cruz. – Rio de Janeiro : Fiocruz, 2020.
33 p.

Elaborado pelo Grupo de Trabalho de Preservação Digital de Acervos da Fiocruz, criado pela Portaria 061/2014-PR de 16/01/2014 e Portaria 405/2017-PR de 23/03/2017.

ISBN 978-65-87465-07-4.

1. Patrimônio. 2. Preservação. 3. Memória (Patrimônio). 4. Brasil. I. Pinheiro, Marcos José de Araújo.

CDD 363.96

SUMÁRIO

Introdução	2
1. Definições	3
2. Objetivo	5
3. Escopo	5
4. Responsabilidades	5
5. Princípios Básicos	6
5.1. Modelo de referência	6
5.2. Formatos de arquivo	7
5.3. Metadados	8
6. Estratégias de Preservação	9
7. Sistemas de Gestão	9
8. Etapas da Preservação	10
8.1 Identificação (Pré-ingestão)	10
8.2. Ingestão	11
8.3. Armazenamento	11
8.4. Acesso	11
9. Gestão de Riscos	11
10. Custos e Financiamento	12
11. Redes de Cooperação	13
12. Atualização do Programa de Preservação Digital	13
Referências	14
Apêndice I – Estrutura Mínima do Plano de Preservação Digital	16

INTRODUÇÃO

Em 2018, a Fundação Oswaldo Cruz (Fiocruz) lançou a [Política de Preservação dos Acervos Científicos e Culturais da Fiocruz](#), no escopo do Preservo – Complexo de Acervos da Fiocruz, com a finalidade de preservar, valorizar e divulgar o patrimônio histórico, cultural e científico institucional e contribuir para a valorização da memória das áreas de Ciências da Saúde e Biomédicas. A Política estabelece os princípios gerais das ações de preservação para os diferentes acervos e coleções científicas e culturais da instituição, incluindo a dimensão da preservação e acesso digital.

O Programa de Preservação Digital de Acervos da Fiocruz (PPD), como parte integrante do Preservo e da [Política de Preservação dos Acervos Científicos e Culturais da Fiocruz](#), estabelece diretrizes para a constituição da infraestrutura física e lógica para o armazenamento de documentos digitais a longo prazo e de um sistema informatizado para preservação dos acervos digitais, bem como orienta a criação de planos de preservação digital específicos para os diferentes acervos.

O presente documento descreve o objetivo, escopo e responsabilidades do PPD, bem como os princípios básicos que norteiam as ações de preservação digital. Além disso, define os padrões e requisitos mínimos para o desenvolvimento dos planos de preservação digital, incluindo a gestão de riscos, a gestão da propriedade intelectual, a seleção de estratégias de preservação, os sistemas de gestão e etapas de preservação a serem adotados, bem como a definição de custos e financiamento.

Orienta-se que esse programa para a preservação digital seja implementado gradualmente em todos os acervos da Fiocruz que possuem documentos digitais, a partir do desenvolvimento de seus próprios planos de preservação digital.

1. DEFINIÇÕES

Acervo digital: um ou mais grupos de documentos digitais com características e/ou fontes que os unificam.

Acesso: direito, oportunidade ou meios de encontrar, recuperar e usar uma informação e/ou documento.

Autenticidade: credibilidade de um documento enquanto documento, isto é, a qualidade de um documento ser o que diz ser e estar livre de adulteração ou qualquer outro tipo de corrupção.

Cópia de segurança: cópia feita com vistas a restaurar os documentos digitais no caso de perda ou destruição do original.

Curadoria: compreende ações voltadas para a seleção, preservação, manutenção, coleção, arquivamento e reavaliação de acervos digitais durante todo o seu ciclo de vida, de forma a garantir a recuperação e acesso a longo prazo da informação armazenada digitalmente.

Digitalização: transformação de informação analógica para formato digital por meio de dispositivo adequado e seguindo requisito de qualidade no seu processo de produtividade.

Direito autoral: abrange os direitos de autor e os direitos conexos. O direito de autor se aplica às obras literárias (como os romances, poemas e peças de teatro), filmes, músicas, obras artísticas (por exemplo, desenhos, pinturas, fotografias e esculturas), aos projetos arquitetônicos e ao software. Os direitos conexos incluem os direitos de artistas executantes sobre suas performances, de produtores de fonogramas sobre suas gravações, e de emissoras de programas de rádio e televisão sobre suas transmissões (Lei nº 9.610, de 19 de fevereiro de 1998 e Lei nº 9.609, de 19 de fevereiro de 1998).

Documento digital: informação registrada ou item do acervo codificada em dígitos binários ou convertida para formato digital acessível e interpretável por meio de sistema computacional.

Emulação: estratégia de preservação digital que se baseia na utilização de recursos computacionais para fazer uma tecnologia atual funcionar com as características de uma obsoleta, aceitando as mesmas entradas e produzindo as mesmas saídas.

Integridade: estado dos documentos que se encontram completos e que não sofreram nenhum tipo de corrupção ou alteração não autorizada nem documentada.

Metadados: dados estruturados que descrevem e permitem encontrar, gerenciar, compreender e/ou preservar documentos digitais ao longo do tempo.

Migração: transferência periódica de um objeto digital de um suporte que está se tornando obsoleto, fisicamente deteriorado ou instável para um suporte mais novo; de um formato obsoleto para um formato mais atual ou padronizado; de uma plataforma computacional em vias de descontinuidade para outra mais moderna.

Nato digital: registro ou documento criado e gerenciado digitalmente.

Nível de preservação: diferentes níveis de qualidade que refletem os requerimentos de armazenamento, integridade, segurança, descrição e acesso dos objetos digitais preservados.

Normalização de formato: conversão de formatos de arquivo para um elenco gerenciável de formatos apropriados para preservação e acesso.

Objeto digital: conjunto de uma ou mais cadeias de bits que registram o conteúdo do objeto e de seus metadados associados. A anatomia do objeto digital é percebida em três níveis: físico, lógico e conceitual.

Open Archival Information System (OAIS): modelo desenvolvido pelo Consultative Committee for Space Data Systems (CCSDS), aprovado para publicação como padrão ISO 14721:2003 e revisado

em 2012, que descreve um esquema conceitual para um repositório digital genérico, aberto a todas as comunidades, com garantias de confiabilidade, que disciplina e orienta um sistema de arquivo dedicado a preservar e manter o acesso à informação digital por longo prazo.

Pacote de Informação: recipiente conceitual definido no modelo OAIS, composto por: informação de conteúdo (objeto de informação a ser preservado e a informação de representação); informação descritiva de preservação (informação necessária para preservar a informação de conteúdo); informação de empacotamento (mantém os componentes do pacote de informação unidos); e informação descritiva (metadados que apoiam a futura localização do objeto de informação por meio de funções de busca e recuperação).

Plano de Preservação Digital: documento resultante da análise de um acervo digital específico, onde são avaliadas e definidas as estratégias de preservação digital adequadas a curto, médio e longo prazo, detalhando tanto os aspectos técnicos como de gestão, incluindo critérios, procedimentos e fluxos de trabalho necessários para garantir a preservação digital.

Preservação digital: conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes digitais, garantindo o acesso e a interpretação de documentos digitais pelo tempo que for necessário.

Propriedade industrial: inclui as patentes de invenção e de modelo de utilidade, as marcas, os desenhos industriais e as indicações geográficas (Lei no 9.279, de 14 de maio de 1996).

Propriedade intelectual: é referente às criações da mente humana: invenções; obras literárias e artísticas; e símbolos, nomes e imagens utilizados em atividades comerciais. É dividida em duas categorias: Propriedade industrial e Direito autoral.

Repositório digital: ambiente de armazenamento e gerenciamento de documentos digitais, formado por elementos de hardware, software e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos.

Repositório digital confiável: repositório digital capaz de manter autênticos os documentos digitais, de preservá-los e prover acesso confiável a eles pelo tempo necessário.

Representante digital: representação em formato digital de um documento originalmente não digital, usado como forma de diferenciá-lo do documento de arquivo criado originalmente em formato digital.

Refreshamento: técnica que consiste em transferir a informação digital de um suporte físico de armazenamento para outro mais atual, antes que o primeiro se deteriore.

Sistema de informação: conjunto organizado de políticas, procedimentos, pessoas, equipamentos e programas computacionais que coleta, produz, processa, armazena, transmite e provê acesso ao acervo.

2. OBJETIVO

Apresentar um conjunto de orientações e procedimentos técnicos relevantes para as ações de preservação dos objetos digitais nos acervos da Fiocruz, no intuito de sistematizar, documentar e padronizar o desenvolvimento de planos de preservação digital específicos para esses acervos.

A preservação digital dos acervos da Fiocruz visa atender à comunidade designada de usuários dos acervos, que inclui pesquisadores, estudantes e profissionais da área de saúde, funcionários da instituição, usuários externos, entre outros.

3. ESCOPO

O Programa de Preservação Digital cobre os acervos científicos e culturais contemplados na Política de Preservação dos Acervos Científicos e Culturais da Fiocruz, compostos pelos acervos arquitetônico, urbanístico e arqueológico, arquivístico, bibliográfico, museológico e as coleções biológicas, e aplica-se tanto a documentos criados originalmente no formato digital (nato digitais), quanto a documentos resultantes da conversão ou digitalização do formato impresso ou analógico para digital (representantes digitais).

4. RESPONSABILIDADES

A implementação do Programa de Preservação Digital de Acervos da Fiocruz cabe às unidades e escritórios responsáveis pela custódia e gestão dos acervos e à Coordenação Geral de Gestão de Tecnologia de Informação (COGETIC), sob a articulação e orientação do Fórum de Preservação Digital, cuja composição e funcionamento é objeto de regulamento específico, com acompanhamento pelo Comitê Gestor do Preservo.

O Fórum de Preservação Digital deverá ser constituído minimamente por representantes da Vice-Presidência responsável pelas ações de informação e comunicação; da Vice-Presidência relacionada às coleções biológicas; da Vice-Presidência responsável pela coordenação das ações de Tecnologia da Informação e Comunicações (TIC) da Fiocruz; além de um responsável pela curadoria do sistema de informação de cada acervo e do seu respectivo profissional de tecnologia da informação.

Como previsto na [Política de Preservação dos Acervos Científicos e Culturais da Fiocruz](#), a preservação dos acervos museológico,

arquitetônico, urbanístico e arqueológico, bem como dos acervos digitais arquivísticos permanentes, estão sob a responsabilidade da Casa de Oswaldo Cruz (COC). O Sistema de Gestão de Documentos e Arquivos (SIGDA) é responsável pelas orientações de preservação dos documentos arquivísticos digitais correntes e intermediários da Fiocruz, de acordo com a Tabela de Temporalidade e Destinação de Documentos da Fiocruz. A preservação do acervo digital bibliográfico, composto pela Rede de Bibliotecas da Fiocruz e pelo Repositório Institucional Arca, está sob responsabilidade do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT). A VideoSaúde – Distribuidora da Fiocruz/Icict fica responsável pela preservação do acervo audiovisual digital sob sua guarda. As ações de preservação digital dos acervos biológicos deverão ser decididas no âmbito da Câmara Técnica de Coleções Biológicas.

A forma de operacionalização do Programa de Preservação Digital varia para cada acervo e deve ser descrita no seu respectivo Plano de Preservação Digital, que deve passar por validação no Fórum de Preservação Digital.

Cada unidade ou escritório responsável pela custódia e gestão de acervos com documentos digitais é responsável, juntamente com a COGETIC, pelo desenvolvimento e execução do Plano de Preservação Digital específico para os seus acervos. A COGETIC deve assumir a responsabilidade compartilhada pela implementação e operação dos aspectos tecnológicos dos sistemas de preservação.

5. PRINCÍPIOS BÁSICOS

De acordo com as principais práticas e modelos internacionais de preservação digital, o Programa de Preservação Digital de Acervos

da Fiocruz tem como compromisso o uso de padrões e protocolos abertos e interoperáveis, a adoção, sempre que possível, de formatos de arquivo abertos e não-proprietários amplamente aceitos, a definição de estruturas padronizadas de metadados descritivos e de preservação, a adesão a práticas de trabalho colaborativo e em rede, bem como a garantia de proteção aos direitos de propriedade intelectual dos criadores. Tais compromissos também devem nortear a escolha das estratégias de preservação e dos requisitos mínimos desejáveis para os sistemas de gestão de preservação.

5.1. Modelo de referência

A Fiocruz tem como objetivo futuro a adesão ao modelo de referência OAIS para a definição da arquitetura de dados e dos aspectos operacionais dos sistemas de gestão da preservação digital.

O modelo OAIS institui as seguintes responsabilidades obrigatórias que deverão ser contempladas no Plano de Preservação Digital de cada acervo, conforme estrutura mínima apresentada no Apêndice I: definição de critérios de priorização para preservação dos objetos digitais; definição dos padrões de formato e nomeação de arquivos; gestão da propriedade intelectual; documentação das políticas e procedimentos de preservação; e acesso ao conteúdo de forma que a comunidade-alvo definida seja capaz de entender e acessar a informação preservada.

A estrutura do Plano de Preservação Digital (Apêndice I) prevê ainda que os acervos cumpram alguns requisitos mínimos em relação ao armazenamento, segurança da informação e gestão dos objetos digitais.

Em uma futura revisão deste Programa, devem ser incluídos os aspectos de auditoria e certificação do modelo OAIS, bem como

recomendações sobre os níveis de preservação digital, a serem desenvolvidos pelo Fórum de Preservação Digital da Fiocruz, com base nos padrões e recomendações internacionais.

5.2. Formatos de arquivo

Para a preservação digital de longo prazo é recomendado o uso de formatos de arquivo abertos e não-proprietários amplamente aceitos, independentes de plataforma e não compactados ou que utilizam compressão sem perda, em especial para as matrizes de preservação. Os formatos de arquivo adotados devem ser compatíveis com o sistema de gestão de preservação digital e devem ser caracterizados e certificados antes da ingestão do objeto digital no sistema.

A produção de representantes digitais de documentos textuais, cartográficos, iconográficos, micrográficos, biológicos e museológicos dos acervos culturais e científicos deve seguir as diretrizes de formato de arquivo estabelecidas no [Manual de Digitalização da Fundação Oswaldo Cruz](#). As especificações de formatos de arquivo para documentos sonoros e audiovisuais e bases de dados, não constantes no Manual de Digitalização, devem seguir as recomendações internacionais de formatos com alta adequação para preservação digital de longo prazo (Quadro 1).

Quadro 1 – Recomendações de formatos com alta adequação para preservação digital de longo prazo de arquivos de texto, imagem, áudio, vídeo e base de dados

TIPO DE ARQUIVO	MATRIZ DIGITAL	DERIVADA DE ACESSO
Texto / Imagem	<ul style="list-style-type: none"> · TIFF (Tagged Image File Format) · PDF/A 	<ul style="list-style-type: none"> · JPEG ou PNG (Portable Network Graphics) · PDF ou PDF/A (múltiplas páginas) · DjVu (grandes formatos)
Áudio	<ul style="list-style-type: none"> · AIFF (Audio Interchange File Format) · ALAC (Apple Lossless Audio Codec) · BWF (Broadcast Wave File) · FLAC (Free Lossless Audio Codec) · WAV (Waveform Audio File Format) 	<ul style="list-style-type: none"> · AAC (Advanced Audio Coding) · MP3 (MPEG-1/2 Audio Layer 3) · OGG (Ogg Vorbis)
Vídeo	<ul style="list-style-type: none"> · AVI (Audio Video Interleave) · MKV/FFV1 (Matroska Multimedia Container) · MOV (QuickTime) 	<ul style="list-style-type: none"> · MP4 (MPEG-4 Parte 14)
Base de dados	<ul style="list-style-type: none"> · JSON (JavaScript Object Notation) · XML (Extensible Markup Language) com schema reconhecido · CSV (Comma-separated value) · SQL (Structured Query Language) 	<ul style="list-style-type: none"> · JSON (JavaScript Object Notation) · XML (Extensible Markup Language) com schema reconhecido · CSV (Comma-separated value) · SQL (Structured Query Language)

Fonte: Fundação Oswaldo Cruz (2018a); Ibict (2017); Brylawski (S.d.); Library of Congress (2019); Iasa (2009)

5. 3. Metadados

Os metadados descritivos dos objetos digitais devem ser baseados no padrão Dublin Core (DC), conforme o conjunto mínimo de metadados e correspondências de campos para as diferentes tipologias de acervos definidos no Manual de Digitalização da Fiocruz.

O conjunto mínimo de metadados de preservação a longo prazo, quando utilizado, deve estar em conformidade com o padrão PREMIS (Preservation Metadata Implementation Strategies) e adaptado para atender aos requisitos específicos de cada acervo.

O uso de metadados técnicos, quando necessário, deve ser definido de acordo com as características e normas específicas de cada tipologia de acervo e devem seguir os padrões e as melhores práticas internacionais.

Recomenda-se que os diferentes metadados associados a um objeto digital sejam reunidos e organizados utilizando um schema XML para estruturação de metadados descritivos, técnicos e de preservação, como o METS (Metadata Encoding and Transmission Standard).

6. ESTRATÉGIAS DE PRESERVAÇÃO

As estratégias específicas de preservação a serem adotadas em cada acervo dependem do formato e da tipologia dos objetos digitais, da infraestrutura e tecnologia existentes, da disponibilidade de recursos humanos e materiais para efetivar o planejamento, e dos riscos identificados, conforme descrito no Plano de Preservação Digital do acervo.

O Plano de Preservação Digital é o documento onde são avaliadas e definidas as estratégias de preservação mais adequadas para o acervo a curto, médio e longo prazo, e que apresenta em detalhes tanto os aspectos técnicos como os de gestão, incluindo critérios, procedimentos, fluxos de trabalho e previsão de custos.

As estratégias de preservação devem considerar soluções consagradas, como a migração e/ou o refrescamento, que contemplem o monitoramento e a atualização de formatos de arquivo

e metadados, de suportes de armazenamento e dos sistemas que os abrigam, a fim de acionar as estratégias correspondentes quando:

- um ou mais formatos de arquivo são considerados sob risco de obsolescência;
- um padrão de metadados é atualizado ou substituído;
- um ou mais suportes de armazenamento estejam sob risco de obsolescência;
- uma versão do sistema é substituída por outra mais recente;
- um sistema é desativado ou substituído por outro.

7. SISTEMAS DE GESTÃO

Uma infraestrutura de repositório de preservação digital será implementada em âmbito institucional pela Fiocruz, a fim de gerenciar a ingestão, armazenamento e acesso a longo prazo do conteúdo e metadados dos seus acervos digitais.

A infraestrutura de repositório de preservação digital deve:

- Dar preferência a sistemas de código aberto e interoperáveis com outros sistemas e padrões;
- Ser baseado nas entidades de informação e entidades funcionais do modelo de referência OAIS;
- Ser minimamente compatível com os padrões de metadados Dublin Core, METS e PREMIS e o protocolo OAI-PMH;
- Oferecer recursos para armazenar e gerenciar as tipologias de conteúdo e metadados definidos pelos acervos da Fiocruz;
- Ter escalabilidade para lidar com volumes crescentes de conteúdo digital que precisam ser ingeridos, verificados, armazenados, preservados e disponibilizados para acesso;
- Oferecer as funcionalidades necessárias para a preservação do conteúdo lógico e intelectual, incluindo: a identificação

e caracterização de artefatos digitais, a validação e manutenção da integridade dos arquivos, a normalização para formatos de preservação, o monitoramento e planejamento da preservação, e o gerenciamento de riscos conhecidos e emergentes;

- Ser independente do sistema de origem da coleção, permitindo restaurar a totalidade dos conteúdos e metadados caso o conteúdo do sistema original seja danificado ou perdido.

8. ETAPAS DA PRESERVAÇÃO

Os tópicos a seguir descrevem conceitualmente o planejamento e as ações necessárias para cada etapa da preservação digital, conforme o Modelo de Referência OAIS. A forma de execução das etapas deve ser descrita no Plano de Preservação Digital, visando o atendimento das recomendações, porém respeitando o grau de desenvolvimento e os recursos disponíveis de cada acervo.

8.1 Identificação (Pré-ingestão)

Antes de efetuar a submissão a um repositório de preservação digital, os produtores do acervo devem preparar pacotes de conteúdo digital, compostos pelo objeto digital acompanhado por seus metadados apropriados. Os curadores já terão selecionado e avaliado o conteúdo destinado à preservação, utilizando critérios próprios, descritos no Plano de Preservação Digital.

Os formatos de arquivo e de metadados enviados devem ser compatíveis com o sistema de gestão de preservação digital. Recomenda-se que os objetos digitais sejam caracterizados e certificados antes de sua ingestão pelo sistema, por meio de uma ferramenta de identificação de formato de arquivo que extraia

também outras informações, como tamanho do arquivo, data da última modificação e localização original. É recomendado ainda o uso de uma ferramenta para geração da soma de verificação (checksum) do arquivo, a fim de possibilitar a aferição de sua integridade em etapas posteriores.

8.2. Ingestão

A ingestão é o processo de transferência do documento digital e de seus metadados correspondentes, estruturados no formato de um Pacote de Informações de Envio ou SIP (Submission Information Package), do produtor para um sistema de gestão de preservação. Durante o processo de ingestão, o sistema executa verificações de qualidade no SIP, o que inclui a certificação dos tipos de arquivo, a validação do seu conteúdo e a normalização dos formatos, conforme necessário. O sistema gera um Pacote de Informações de Arquivamento ou AIP (Archival Information Package), que inclui a criação ou aprimoramento de metadados associados ao documento digital, incluindo os metadados de preservação, quando aplicável, e procede com a transferência do AIP para um repositório de armazenamento de longo prazo.

8.3. Armazenamento

Esta etapa trata do armazenamento, manutenção e recuperação dos AIPs, que são enviados para armazenamento permanente de acordo com diferentes critérios (formatos, taxas de uso esperadas etc.) configurados no sistema de gestão de preservação. O armazenamento de arquivos exige infraestrutura técnica especializada e suporte para a duplicação do conteúdo digital em sistemas locais e geograficamente distribuídos. A manutenção dos

AIPs inclui a migração para novos formatos, conforme necessário, a verificação de erros de arquivos, a implementação de estratégias para recuperação de desastres e o fornecimento de cópias de acesso dos documentos digitais aos usuários.

8.4. Acesso

A etapa de acesso permite aos usuários (também chamados “consumidores” no modelo OAIS) encontrar informações sobre o documento digital em um repositório e acessar o seu conteúdo. O sistema de gestão de preservação gera um Pacote de Informações de Divulgação ou DIP (Dissemination Information Package) em resposta a uma solicitação do usuário e o disponibiliza por meio de uma interface de usuário para acesso aos arquivos. Um DIP compreende um derivado do(s) AIP(s) apropriado(s) contendo os arquivos armazenados e seus metadados relevantes.

9. GESTÃO DE RISCOS

A Gestão de Riscos dos acervos digitais da Fiocruz deve ter como referências fundamentais duas normas: a ISO 31000, que estabelece princípios e orientações sobre a gestão de riscos; e a ISO/IEC 31010, que estabelece técnicas e define ferramentas que apoiam e dão estrutura à avaliação de riscos.

A Gestão de Riscos deve ser estruturada de acordo com as etapas consecutivas do processo de gerenciamento descrito na norma técnica ISO 31000, que incluem: 1) Estabelecer contexto (interno e externo); 2) Identificar os riscos; 3) Analisar os riscos; 4) Avaliar e priorizar os riscos; 5) Reduzir os riscos segundo as prioridades estabelecidas. Além disso, todo esse processo deve ser

periodicamente monitorado e revisado.

O processo de gerenciamento de riscos deve ser pautado pela [Política de Gestão de Integridade, Riscos e Controles Internos da Gestão da Fiocruz](#) e pelo [Guia de Gestão de Riscos da Fiocruz](#) e deve observar a legislação relacionada à propriedade intelectual.

O gerenciamento de riscos dos acervos digitais será parte do escopo dos processos de gestão de riscos e gestão da continuidade de negócios conduzidos pela COGETIC.

10. CUSTOS E FINANCIAMENTO

É necessário analisar, controlar e gerenciar, de maneira ativa, os custos envolvidos na preservação a longo prazo de documentos digitais. O responsável pela curadoria do sistema de informação de cada acervo e o profissional de tecnologia da informação da unidade, com o apoio dos profissionais da COGETIC, deverão estimar os custos da preservação digital do acervo e documentar as conclusões no respectivo Plano de Preservação Digital, podendo incluir os seguintes indicadores:

- Custos de implantação e manutenção do sistema de gestão de preservação, incluindo espaço físico, equipamentos (servidores, mídias de armazenamento e backup) e softwares (sistemas operacionais, antivírus, ferramentas de preservação etc.);
- Custos de pessoal, com base nas exigências de qualificação e carga horária;
- Custos operacionais das ações de preservação, como digitalização, controle de qualidade, conversão, migração de formatos e suportes, de acordo com as estratégias de preservação elegidas;

- Custos relativos à gestão de direitos de propriedade intelectual, como assessoria jurídica, realização de inventário de direitos, obtenção de assinaturas em instrumentos jurídicos de cessão ou autorização, buscas diligentes por autores e/ou titulares de direitos, monitoramento da utilização pelos usuários finais, utilização de estratégias para evitar a cópia ou o mau uso por terceiros não autorizados e, finalmente, realização de avaliação de risco para digitalização e disponibilização de obras órfãs e custos judiciais em virtude da digitalização e disponibilização não autorizada de conteúdos pela própria Fiocruz ou por terceiros;
- Outros custos relacionados, incluindo fatores indiretos como a periodicidade das ações de preservação e monitoramento, integração com outros sistemas, previsão de crescimento do volume do acervo, atendimento a demandas técnicas e/ou legais etc.

Sugere-se que o Plano de Preservação Digital seja pautado por um modelo de custo, instrumento que permite obter uma melhor compreensão e um controle mais efetivo sobre como os custos são distribuídos, a fim de subsidiar decisões futuras relativas ao financiamento das atividades de preservação digital.

O financiamento das ações ordinárias e estratégicas para a preservação digital dos acervos pela Fiocruz, constituindo o Tesouro Nacional sua principal fonte de recursos, é feito também com recursos provenientes de outras fontes, que incluem principalmente aqueles captados por meio de editais de instituições públicas e privadas de fomento, leis de incentivo fiscal, patrocínio e subvenções.

11. REDES DE COOPERAÇÃO

Tendo como princípio a valorização da colaboração e do trabalho em rede, a Fiocruz compromete-se a estabelecer diálogo e parcerias com instituições nacionais e internacionais que possuam os conhecimentos necessários na área de preservação digital, e prevê a participação em redes de cooperação para preservação distribuída.

Através de projetos estratégicos de colaboração, a Fiocruz deseja garantir uma posição-chave nas redes internacionais que trabalham com a preservação digital, de forma a posicionar-se, legalmente, politicamente e em termos de conhecimento, métodos e ferramentas, em par com instituições comparáveis no mundo.

12. ATUALIZAÇÃO DO PROGRAMA DE PRESERVAÇÃO DIGITAL

A fim de garantir que as orientações descritas atendam às melhores práticas, ao estado atual dos acervos e aos contínuos desenvolvimentos tecnológicos, a revisão e atualização do Programa de Preservação Digital de Acervos da Fiocruz deverão ser realizados em intervalos máximos de quatro anos, sob a orientação do Fórum de Preservação Digital.

Além da revisão periódica, outros eventos podem demandar a revisão extemporânea do Programa, como a entrada de novos acervos, a mudança nos objetivos da preservação ou nas instâncias institucionais responsáveis, a adoção de novas estratégias tecnológicas ou a identificação de riscos iminentes, ficando a critério do Fórum de Preservação Digital a convocação de seus membros para discussão.

Este documento atualizado deve estar publicamente disponível no site da Fiocruz, com o objetivo de apoiar a colaboração nacional e internacional em preservação digital.

REFERÊNCIAS

Documentos internos

FUNDAÇÃO OSWALDO CRUZ. Manual de digitalização. Rio de Janeiro: Fiocruz, 2018a.

FUNDAÇÃO OSWALDO CRUZ. Política de preservação dos acervos científicos e culturais da Fiocruz. Rio de Janeiro: Fiocruz, 2018b.

Documentos consultados

BARBEDO, Francisco (Coord.). Recomendações para a produção de Planos de Preservação Digital. Lisboa: Direção Geral de Arquivos, 2011.

BECKER, Christoph et al. Distributed preservation services: Integrating planning and actions. In: CHRISTENSEN-DALSGAARD, B. et al. (Ed.) Research and Advanced Technology for Digital Libraries. Berlin: Springer, 2008. p. 25-36.

BRYLAWSKI, Sam et al. (Ed.) ARSC Guide to Audio Preservation. Eugene, OR: Association for Recorded Sound Collections,[2015].

CAPLAN, Priscilla. Understanding PREMIS. [Washington, DC]: Library of Congress, 2017.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). Câmara Técnica de documentos eletrônicos. Diretrizes para a implementação de repositórios digitais confiáveis de documentos – RDC-Arq. Rio de Janeiro: Arquivo Nacional, 2015. Disponível em: http://www.conarq.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf. Acesso em: 15 out. 2020.

HUGHES, Lorna M.; GREEN, David. Digitizing collections: strategic issues for the information manager. London: Facet Publishing, 2004.

IASA TECHNICAL COMMITTEE. Guidelines on the Production and Preservation of Digital Audio Objects. London, 2009.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Rede Cariniana. Lista de formatos preservados. Brasília, DF, 2017. Disponível em: <http://cariniana.ibict.br/index.php/listas/lista-de-formatos-preservaveis>. Acesso em: 15 jun. 2017.

LAVOIE, Brian; GARTNER, Richard. Preservation Metadata. 2. ed. [York]: Digital Preservation Coalition, 2013. [DPC Technology Watch Report.

LIBRARY OF CONGRESS. Recommended Formats Statement 2019-2020. Washington, DC, [2019]. Disponível em: <https://www.loc.gov/preservation/resources/rfs/TOC.html>. Acesso em: 19 jul. 2019.

NATIONAL RECORDS OF SCOTLAND. The National Records of Scotland and born digital records – a strategy for today and tomorrow. Edimburgo, 2014.

SAYÃO, Luís Fernando. Uma outra face dos metadados: informações para a gestão da preservação digital. Encontros Bibli, Florianópolis, v. 15, n. 30, p. 1-31, out. 2010.

SIZE, Any; BROWN, Adrian. Practical digital preservation: a how-to guide for organizations of Any Size. London: Facet Publishing, 2013.

SLATS, Jacqueline; VERDEGEM, Remco. Cost Model for Digital Preservation. [S.l.: s.n.], [2005]. Disponível em: https://dlmforum.typepad.com/Paper_RemcoVerdegem_and_JS_CostModelfordigitalpreservation.pdf. Acesso em: 15 jul. 2020.

STRODL, Stephan; RAUBER, Andreas. Preservation planning in the OAIS model. Data Analysis and Knowledge Discovery, v. 24, n. 1, p. 61-68, 2008.

THE BRITISH LIBRARY. Sustaining the value: The British Library digital preservation strategy 2017-2020. London, [2016]. Jan. 2017.

THE ROYAL LIBRARY OF DENMARK. Strategy for long-term preservation of digital collection materials at the Royal Library. Copenhagen, 2014.

APÊNDICE I – ESTRUTURA MÍNIMA DO PLANO DE PRESERVAÇÃO DIGITAL

1. Introdução

Apresentação do Plano de Preservação Digital.

2. Cenário institucional

2.1. Estrutura organizacional: Descrição da unidade, mandato, posição na estrutura institucional.

2.2. Políticas e normas relacionadas: Mencionar o PRESERVO e qualquer outra política interna da unidade ou do campo de atuação relacionada a preservação digital do acervo. Listar também legislações, normas ou outros dispositivos legais a que se submete a gestão do acervo digital, se aplicável.

3. Descrição do acervo

3.1. Identificação e escopo: Descrição do acervo digital a ser preservado, incluindo informações sobre os objetos digitais, seus metadados e características relevantes.

Exemplos de perguntas a serem respondidas:

- Qual é a missão e objetivos do acervo?
- Quem é o público-alvo do acervo?
- Qual o número total aproximado de objetos digitais no acervo?
- Quais são as tipologias de documentos presentes no acervo?
- Quais são formatos de arquivo adotados?
- Os objetos digitais possuem metadados descritivos?

- Os metadados utilizados seguem algum padrão? Se sim, qual? Se não, há planos para a adoção de algum padrão de metadados?
- Existe um conjunto mínimo de metadados para a inclusão de objetos digitais no acervo?
- Os objetos digitais do acervo seguem um padrão para nomenclatura de arquivos?

3.2. Aspectos legais: Dentre os aspectos a serem considerados na seleção de itens do acervo a serem digitalizados, as questões legais merecem destaque.

Aqui serão brevemente destacados alguns pontos a serem observados, mas é fortemente recomendado o auxílio de uma assessoria jurídica especializada.

Exemplos de perguntas a serem respondidas:

- O item ou conjunto de itens a serem digitalizados são protegidos pelo direito autoral?
- O inventário de direitos patrimoniais sobre as obras dos acervos a serem digitalizados está atualizado?
- No caso de fotografias e obras audiovisuais, a Fiocruz possui as devidas autorizações de uso de nome e imagem assinadas por todos os indivíduos retratados?
- A Fiocruz possui estratégias para monitorar o acesso e proteger os conteúdos digitalizados da utilização ou cópia indevida por parte de terceiros não autorizados?

3.3. Aspectos técnicos: Descrição da infraestrutura tecnológica, incluindo informações sobre o sistema no qual os objetos digitais e metadados estão armazenados e quaisquer restrições técnicas.

Exemplos de perguntas a serem respondidas:

- A unidade conta com um repositório digital ou base de dados para organização dos objetos digitais do acervo? Se sim, qual?
- Onde está localizado fisicamente o sistema que armazena os objetos digitais?
- Em quais tipos de dispositivo (servidor, mídia externa, etc.) os objetos digitais do acervo estão armazenados?
- Qual é o tamanho total aproximado do acervo (em GB)?
- Qual a previsão de crescimento anual do acervo digital (em GB)?
- Os objetos digitais do acervo requerem hardware ou software específico para interpretar e apresentar o arquivo?
- Existe alguma mudança tecnológica prevista em relação ao armazenamento e gestão dos objetos digitais do acervo no futuro próximo?

3.4. Aspectos organizacionais: Recursos humanos e financeiros alocados e/ou disponíveis para a gestão dos objetos digitais do acervo, incluindo quaisquer restrições administrativas e/ou orçamentárias.

Exemplos de perguntas a serem respondidas:

- Quais recursos a unidade tem disponíveis para a preservação digital do acervo (incluindo pessoal, equipamentos, sistemas, orçamento)?
- Existe algum planejamento para o investimento em hardware, softwares, mídias, equipamentos ou formação de pessoal para a preservação digital?
- Existe um fluxo de trabalho estabelecido para a gestão dos objetos digitais do acervo? O fluxo de trabalho está documentado?

3.5. Práticas correntes de preservação: listar as atuais práticas de preservação adotadas no acervo, se aplicável.

Exemplos de ações relacionadas à preservação digital:

- Realização de cópias de segurança regularmente com função de restauração;
- Inspeção regular das cópias de segurança para verificar a qualidade e acesso aos dados;
- Armazenamento de cópias de segurança em diversos locais geograficamente distribuídos;
- Utilização de mais de um tipo de armazenamento físico para proteção contra falhas de formato;
- Ambiente de armazenamento em condições estáveis de temperatura e umidade relativa;
- Manutenção de versões anteriores de hardware, sistema operacional e softwares necessários para ler todos os formatos de arquivo no acervo digital;
- Monitoramento do processo de obsolescência de formatos de arquivo, suportes e mídias de arquivos;
- Migração de formatos de arquivo, suportes e mídias sob risco de obsolescência;
- Criação de valores de soma de verificação (checksum) para os objetos digitais na ingestão de arquivos;
- Verificação de integridade do checksum dos objetos digitais no momento da ingestão ou a intervalos regulares;
- Controle e/ou registro de quem tem acesso para movimentar, apagar ou alterar os arquivos digitais.

4. Avaliação e gestão de riscos

Elencar possíveis ameaças internas e externas ao suporte físico e lógico do acervo. Para cada item, incluir a descrição do risco, probabilidade de perda de informação, possíveis consequências, eventuais controles existentes para mitigação de risco e o nível do risco.

Exemplos de perguntas a serem respondidas:

- Como e com que periodicidade é feita a verificação de vírus e outras ameaças nos arquivos digitais?
- Há algum procedimento de quarentena para os objetos digitais que são incorporados ao acervo?
- São geradas cópias de segurança (backup) do acervo? Quantas cópias são feitas e com que periodicidade? Onde são armazenados os arquivos de backup? A qualidade do backup é testada?
- Os objetos digitais (e/ou o sistema onde se encontram) estão armazenados em local considerado fisicamente seguro?
- O acervo possui um inventário com o nome e a localização de todos os objetos digitais que permita recuperar todos os objetos digitais?
- Os objetos digitais do acervo podem ser substituídos em caso de perda?
- Qual é o provável impacto para a instituição caso um objeto digital deste acervo seja perdido e não seja recuperável?
- O sistema possui documentação para manutenção, gestão e recuperação de dados?
- Existe documentação sobre o gerenciamento das mídias de armazenamento, incluindo a sua atualização, manutenção e procedimentos de backup?
- Existe um plano de contingência para restauração do sistema ou dos objetos digitais?

5. Planejamento da preservação

5.1. Metas de preservação: Definir volume ou recorte da coleção que deve ser preservada, critérios de seleção dos objetos digitais, ordem de prioridade da preservação e

prazos previstos para que isso ocorra.

5.2. Estratégias de preservação avaliadas: Considerar as soluções potenciais para preservação do acervo, de acordo com as metas de preservação, incluindo o planejamento para evitar a obsolescência de formatos, metadados, suportes e sistemas, a curto, médio e longo prazo. Sempre que possível, realizar testes de cada alternativa com uma amostra do acervo e relatar os resultados, detalhando tanto os aspectos técnicos como de gestão.

Para cada alternativa proposta, sugere-se a descrição de:

- Fluxo de trabalho, incluindo diagramas para preparação, ingestão, gerenciamento e acesso aos objetos digitais e metadados;
- Proposta de tecnologia (hardware e software);
- Recursos humanos necessários;
- Resumo do custo para implementação do sistema;
- Resultados do teste (ou relato de evidências de soluções que não foram avaliadas);
- Discussão da solução proposta, incluindo custo, benefícios, impacto, riscos e oportunidades, e em que grau a solução proposta atende aos requisitos e restrições.

5.3. Identificação e justificativa da estratégia selecionada: Identificar a estratégia selecionada, explicitando o motivo da escolha.

6. Ações de preservação

Nessa seção, são descritos os critérios, procedimentos e fluxos operacionais necessários para a preservação digital do acervo, de

acordo com a estratégia de preservação selecionada.

As seguintes responsabilidades obrigatórias deverão ser contempladas no Plano de Preservação Digital, conforme previsto no modelo OAIS:

- Definição de critérios de seleção e priorização para preservação dos objetos digitais
- Definição e controle dos padrões de formato de arquivo;
- Definição e controle da nomenclatura dos objetos digitais;
- Gestão da propriedade intelectual sobre o conteúdo preservado;
- Documentação das políticas e procedimentos de preservação;
- Descrição das formas de acesso à informação preservada pela comunidade-alvo.

Além disso, os seguintes itens são considerados obrigatórios para a aprovação do Plano de Preservação Digital:

- Garantia de duplicação do armazenamento dos arquivos digitais em locais distintos e geograficamente dispersos;
- Verificação sistemática de vírus; Inventário dos arquivos digitais;
- Fluxogramas de trabalho detalhando todas as tarefas necessárias para realizar as ações de preservação, indicando a periodicidade e o responsável por cada tarefa;
- Adoção de vocabulário controlado para descrição de assuntos do acervo, sendo recomendado o DeCS (Descritores em Ciências da Saúde) como vocabulário base. Para assuntos específicos não cobertos pelo DeCS, deve ser adotado vocabulário controlado específico. Em última instância, prever a criação de descritores.

O Fórum de Preservação Digital recomenda ainda a inclusão das seguintes informações:

- Descrição dos padrões e do conjunto mínimo de metadados descritivos, técnicos e de preservação dos objetos digitais

- do acervo;
- Descrição das ações ou ferramentas utilizadas para garantir a integridade e autenticidade dos objetos digitais do acervo;
 - Plano de sucessão contendo as ações necessárias para a restauração dos objetos digitais e seus metadados, em caso de perda do acervo;
 - Previsão de análise de vulnerabilidade dos softwares de acesso e aplicação de medidas para mitigar os riscos.

7. REVISÃO DO PLANO DE PRESERVAÇÃO

Indicar a periodicidade de revisão do Plano de Preservação Digital e determinar eventos futuros que possam desencadear a revisão extemporânea do documento.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz