



FUNDAÇÃO OSWALDO CRUZ
PLANO DE PRESERVAÇÃO DIGITAL
COLEÇÃO MAIS MÉDICOS

Fundação Oswaldo Cruz
Casa de Oswaldo Cruz
Departamento de Arquivo e Documentação

Plano de Preservação Digital Coleção Mais Médicos

Rio de Janeiro
2022



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 os 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:

Casa de Oswaldo Cruz

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

Site: www.coc.fiocruz.br

Plano de Preservação Digital – Coleção Programa Mais Médicos

Versão: 1.1

Data de atualização: 1 de outubro de 2021

Autores: Aline Lopes de Lacerda, Felipe Almeida Vieira, Marco Dreer Buarque

Colaboração: Jeferson Mendonça dos Santos Silva, Roberto Jesus Oscar e Vinícius Pequeno de Souza

Edição de texto: Regina Celie Simões Marques

Capa, projeto gráfico e diagramação: Silmara Mansur

Ficha Catalográfica

F981p Fundação Oswaldo Cruz. Casa de Oswaldo Cruz.

Plano de preservação digital : coleção Mais Médicos / Casa de Oswaldo Cruz.
Departamento de Arquivo e Documentação. – Rio de Janeiro: Fiocruz/COC, 2022.
Edição digital.

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

1. Patrimônio Cultural. 2. Preservação. 3. Organização do Conhecimento. I.
Lacerda, Aline Lopes de. II. Buarque, Marco Dreer. III. Vieira, Felipe Almeida.

SUMÁRIO

1.	Introdução	4
2.	Cenário institucional	5
2.1.	Estrutura organizacional	5
2.2.	Políticas de preservação relacionadas	7
2.3.	Aspectos legais	9
3.	Descrição do acervo	10
3.1.	Identificação e escopo	10
3.2.	Aspectos organizacionais	13
3.3.	Aspectos técnicos	15
3.4.	Requisitos dos usuários	16
3.5.	Práticas correntes de preservação	17
4.	Avaliação de riscos	18
5.	Planejamento de preservação	24
5.1.	Prioridades de preservação	24
5.2.	Estratégia de preservação avaliada	24
5.3.	Funções e responsabilidades	26
5.4.	Custos e fontes de financiamento	27
6.	Ações de preservação	27
7.	Revisão do Plano de Preservação	29
	APÊNDICE A – Instruções de uso das ferramentas gratuitas	31
	APÊNDICE B – Permissão de acesso e trilha de auditoria	35

1. INTRODUÇÃO

O presente documento tem por finalidade indicar estratégias e ações de preservação digital das fotografias nato digitais¹ da coleção fotográfica Programa Mais Médicos (coleção MM). A coleção foi anteriormente identificada, classificada e descrita na [Base Arch](#), onde se encontra acessível à consulta on-line. Contudo é necessário estabelecer padrões de conduta para a salvaguarda desse material visando sua longa permanência e acesso. Este é o objetivo do presente documento.

Este Plano de Preservação Digital vincula-se ao Programa de Preservação Digital de Acervos da Fiocruz, documento que estabelece diretrizes para constituição de infraestrutura física e lógica para armazenamento de documentos digitais a longo prazo e de sistema informatizado para preservação dos acervos digitais. As orientações e diretrizes constantes em ambos os documentos visam assegurar a preservação dos atributos de identidade e integridade dos documentos de arquivo, bem como da presunção de autenticidade dos mesmos, que serão reforçadas pela identificação e registro dos contextos de produção e da relação orgânica entre os documentos de um fundo ou coleção, informações que deles jamais devem ser apartadas.

A Coleção Programa Mais Médicos foi identificada como o conjunto arquivístico apropriado para a elaboração deste Plano de Preservação Digital por ser composto inteiramente de imagens nato digitais e já submetido a tratamento arquivístico. Entende-se que o formato nato digital das imagens que compõem a coleção, seu valor como acervo histórico e o status de acervo organizado reúnem os principais requisitos para a consolidação de processos

1. Fotografia nato digital é uma imagem obtida por processo eletrônico, através do sensor interno das câmeras digitais.

e ações voltadas para a preservação de documentos arquivísticos desse gênero e formato, em idade permanente.

Espera-se, portanto, que este plano sirva de referência para outros da mesma natureza, a saber, voltados para a preservação de documentos de formato nato digital, de caráter permanente e já organizados e descritos arquivisticamente.

2. CENÁRIO INSTITUCIONAL

2.1. Estrutura organizacional

A Casa de Oswaldo Cruz (COC) é uma unidade da Fundação Oswaldo Cruz. Criada em 1986, configura-se como um centro de pesquisa, documentação e educação em história, memória, preservação do patrimônio cultural e divulgação em ciências e saúde.

A COC é responsável pela preservação e pelo acesso a um diversificado patrimônio da Fiocruz composto por edificações arquitetônicas e acervos documental, bibliográfico e museológico. Alguns elementos do patrimônio arquitetônico, arqueológico e urbanístico são tombados nos níveis federal, estadual ou municipal, a exemplo de edificações ecléticas e modernistas no campus Manguinhos e sítios arqueológicos ali existentes.

Sob a responsabilidade do Departamento de Arquivo e Documentação (DAD) da COC, o acervo arquivístico histórico possui hoje cerca de 140 conjuntos documentais, entre fundos arquivísticos de unidades da Fiocruz e de instituições privadas da área da saúde já extintas, além de arquivos pessoais e coleções documentais de cientistas, médicos e sanitaristas com atuação nos campos das ciências biomédicas, sociais e da saúde. Três

desses conjuntos são reconhecidos pelo Programa Memória do Mundo da Unesco.² Nesse acervo destacam-se os gêneros textual, com cerca de 1,5 km de extensão, e fotográfico, que atualmente ultrapassa a marca dos 150 mil itens.

Duas bibliotecas são responsáveis pelo acervo bibliográfico da COC: a Biblioteca de História das Ciências e da Saúde, com mais de 80 mil obras de história da medicina, história da saúde pública, história, sociologia e filosofia da ciência, do século 17 aos dias atuais; e a Biblioteca de Educação e Divulgação Científica, vinculada ao Museu da Vida, que contempla as áreas de divulgação científica, educação, museologia, ciências da vida, saúde e literatura infanto-juvenil.

O acervo museológico da COC encontra-se sob a guarda da Reserva Técnica Museológica do Museu da Vida e possui mais de três mil itens, com ênfase na história institucional. Abrange o período compreendido entre meados do século XIX e a atualidade e é composto de instrumentos e equipamentos de laboratório, materiais e maquinário utilizados na produção de medicamentos e vacinas, instrumentos médicos, mobiliário, indumentária, objetos pessoais de cientistas da instituição e uma pinacoteca.

A maior parte do patrimônio histórico-cultural da Fiocruz sob responsabilidade da COC está à disposição do público por meio de atendimento para consulta aos acervos arquivístico e bibliográfico, exposições dos itens museológicos e sistemas de consulta on-line na Web.

2. Os arquivos dos cientistas Oswaldo Cruz e Carlos Chagas, bem como os negativos de vidro do fundo Instituto Oswaldo Cruz foram reconhecidos como patrimônio documental da humanidade pelo Programa Memória do Mundo da Unesco. Criado em 1992, o Programa tem por objetivos assegurar a preservação do patrimônio documental com significação mundial, regional ou nacional, assegurar o acesso a ele e disseminar o conhecimento sobre esse patrimônio.

A COC dispõe de quadro profissional qualificado, com pesquisadores que atuam como líderes de pesquisa no CNPq em grupos voltados ao estudo histórico da medicina, das ciências, das políticas e das tecnologias. Realiza também pesquisa na área de arquivologia, preservação de acervos e divulgação e popularização da ciência. Desenvolve atividades de educação formal e não formal e possui três programas de pós-graduação: História das Ciências e da Saúde (Mestrado e Doutorado); Mestrado acadêmico em Divulgação da Ciência, Tecnologia e Saúde; Mestrado Profissional em Preservação e Gestão do Patrimônio Cultural das Ciências de Saúde, além do Curso de Especialização em Divulgação e Popularização da Ciência. Oferece, ainda, cursos de formação inicial para produtores culturais e graduandos interessados nas áreas de divulgação e popularização da ciência, conservação de bens culturais e gestão da documentação, além de diversos cursos livres. As ações de educação envolvem mais de 200 alunos anualmente.

O Museu da Vida é o departamento da COC dedicado à divulgação e popularização da ciência. Seu principal espaço de visitação no campus de Manguinhos, aberto e gratuito, oferece ao público circuitos interativos com atividades voltadas à reflexão sobre as ciências e a saúde e suas relações com a vida, a sociedade e a história. O Museu da Vida desenvolve também atividades extramuros com o Caminhão da Ciência, que disponibiliza atividades e exposições a populações de municípios do Sudeste e a comunidades de periferia e favelas do Rio de Janeiro.

2.2. Políticas de preservação relacionadas

Ao custodiar acervos histórico-culturais das ciências e da saúde, a COC assume a responsabilidade pela salvaguarda e difusão de bens que integram o patrimônio científico-cultural da Fiocruz,

amplamente reconhecido por agências públicas e pela sociedade. No cumprimento de tal atribuição, colabora com a instituição na organização e implantação de políticas, estruturas e processos relativos a esses acervos buscando aliar sua preservação à ampliação de seu uso social e geração de pesquisas nos campos das ciências e da saúde.

Em sintonia com conceitos e práticas nacionais e internacionais que têm como princípios centrais a conservação preventiva, o gerenciamento de riscos, a conservação integrada e a preservação sustentável, a COC formulou em 2013 a Política de Preservação e Gestão de Acervos Culturais das Ciências e da Saúde.³ O documento, inovador ao estabelecer diretrizes e princípios a acervos diversos (como arquitetônico, arquivístico, bibliográfico e museológico), também afirma a pesquisa e a educação como fundamentais e estratégicas para a estruturação de políticas de preservação.

No âmbito da preservação de documentos digitais, a presença crescente de documentos nato digitais em acervos arquivísticos históricos recebidos em doação tem levado a COC a desenvolver estratégias de longo prazo, amparadas em diálogo com políticas institucionais, em especial na Política de Preservação dos Acervos Científicos e Culturais⁴ e na Política de Acesso Aberto ao Conhecimento da Fiocruz⁵. Tais políticas estabelecem princípios e diretrizes para a preservação do patrimônio científico e cultural da instituição e o acesso aberto ao conhecimento dele derivado.

3. Disponível em: http://www.coc.fiocruz.br/images/PDF/politica_preservacao_gestao_acervos_coc.pdf; acesso em 23 mar. 2022.

4. Disponível em: <https://portal.fiocruz.br/documento/politica-de-preservacao-dos-a-cervos-cientificos-e-culturais-da-fiocruz>; acesso em 23 mar. 2022.

5. Disponível em: https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria_-_politica_de_acesso_aberto_ao_conhecimento_na_fiocruz.pdf; acesso em 23 mar. 2022.

Internamente, o tema dos documentos digitais tem sido central em projetos de integração de acervos culturais da Fiocruz – como o projeto PRESERVO – e a elaboração do Programa de Preservação Digital, do qual este documento deriva.

2.3. Aspectos legais

A coleção Programa Mais Médicos foi doada à Fiocruz por seu autor/produtor, o fotógrafo Araquém Alcântara, por meio da assinatura de Termo de Doação. Pelo termo estabeleceu-se a incorporação do conjunto documental ao patrimônio científico-cultural da Fiocruz, que passou então a assumir o domínio, a posse e o direito de utilizá-lo e disponibilizá-lo para estudos, pesquisas e produtos de caráter acadêmico e cultural. O mesmo Termo esclarece que cabe à Fiocruz, por meio da COC, a obrigação de organizar, inventariar e conservar a coleção para fins de pesquisa e viabilizar seu acesso.

Ao incorporar acervos dessa natureza a seu patrimônio e uma vez cedidos os direitos de uso das imagens que os compõem, a instituição fica também autorizada a dar acesso ampliado a eles, conforme condições determinadas pela COC, a observância da legislação vigente sobre arquivos e o que determinam a Constituição e o Código Civil com relação aos direitos de proteção à privacidade, intimidade, vida privada, honra e imagem dos indivíduos.

3. DESCRIÇÃO DO ACERVO

3.1. Identificação e escopo

A Coleção Programa Mais Médicos originou-se do trabalho fotográfico de Araquém Alcântara Pereira, realizado entre março e outubro de 2015. O acervo, em formato digital e composto por 673 arquivos de imagens, foi transferido para a COC em novembro de 2016, quando os direitos de uso das imagens também foram cedidos à instituição.

Araquém Alcântara utilizou 86 fotografias no livro “Mais Médicos”, de sua autoria.⁶ A obra narra por meio de imagens a rotina de 43 profissionais de saúde, brasileiros e estrangeiros, que atuaram entre março e outubro de 2015 no Programa Mais Médicos em 36 cidades de diferentes regiões do país, suas relações com os pacientes, as dificuldades de acesso por eles enfrentadas, bem como aspectos das localidades e condições de vida da população assistida.

O Programa Mais Médicos foi lançado pelo governo federal em 8 de julho de 2013 por medida provisória convertida na lei 12.871, de 22 de outubro de 2013. Em sua origem, teve como principal objetivo diminuir a carência de médicos nas regiões consideradas prioritárias para o SUS, devido à maior vulnerabilidade de seus habitantes e à dificuldade para atrair e ali fixar profissionais. Para sua implantação recorreu-se a provimento emergencial e recrutamento de profissionais brasileiros e estrangeiros, graduados no Brasil e exterior.

Em 2015 faziam parte do Programa 18.240 médicos incluindo profissionais procedentes de mais de 40 países, a exemplo de médicos cubanos que participavam da cooperação com a

6. ALCÂNTARA, Araquém. Mais Médicos. São Paulo: Terrabrasil, 2015.

Organização Pan-Americana da Saúde (Opas). Naquele ano 4.058 municípios foram atendidos pelo Programa contemplando 34 distritos sanitários especiais de saúde indígena, unidades básicas de saúde fluviais de atendimento da população ribeirinha na Amazônia, como também comunidades rurais do semi árido nordestino, quilombolas, assentamentos e periferias e morros de grandes cidades.

A organização da coleção pelo DAD deu-se entre os meses de fevereiro e maio de 2019 e constou das etapas de identificação, organização, descrição, inserção na [Base Arch](#) – repositório de informações sobre os arquivos permanentes da Fiocruz – e revisão.

Ao darem entrada no DAD, os documentos encontravam-se armazenados em duas unidades de disco rígido externo, organizados em quatro pastas e totalizando 410 arquivos em formato CR2 (formato do tipo RAW⁷ proprietário originário da câmera fotográfica), 250 em formato TIFF e 13 em formato JPEG. A identificação das imagens tomou como ponto de partida a mencionada publicação de autoria de Araquém Alcântara, que forneceu informações sobre localidades, pessoas e entidades fotografadas. Nessa etapa recorreu-se também aos metadados de data e hora originais das fotos para distinção entre originais e cópias, como também para a identificação dos conjuntos de fotografias produzidos em sequência e em determinada localidade. Ainda com base nos metadados dos arquivos originais, foi possível identificar equipamento e softwares utilizados na produção e edição das fotos: uma câmera DSLR modelo Canon EOS-1D X, que gera arquivos RAW(formato CR2); e, a partir

7. De acordo com o Conarq, RAW “é o formato de arquivo digital das câmeras digitais e de escaners. São formatos proprietários. Em geral as imagens em RAW são convertidas para formatos como o TIFF ou JPEG. Em fotografia digital são muitas vezes denominados de ‘negativos digitais’ pois são o registro em arquivo digital”. Disponível em: https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/Recomendacoes_digitalizacao_completa.pdf ; p.9, nota 16; acesso em 25 jun. 2020.

destes, o software de edição Adobe Photoshop Lightroom 5.7, que gera arquivos derivados em formato TIFF ou JPEG.

Com as informações obtidas na etapa de identificação, para fins arquivísticos consideraram-se itens documentais da coleção os arquivos de imagens únicos e originalmente produzidos pela câmera fotográfica (em formato CR2) e aqueles produzidos e editados a partir deles (em formato TIFF) por Araquém Alcântara, cujas características de conteúdo são significativamente distintas dos primeiros em aspectos variados, entre eles o padrão de cor. Tal definição teve como diretriz a lógica de produção dos documentos, decorrente da própria atividade do fotógrafo e expressa na publicação mencionada. Estabeleceu-se assim que a coleção é formada por 634 itens originais, sendo 391 imagens CR2 e 243 imagens TIFF, no total de 28,3 gigabytes (GB) de informação. Por decorrência, foram desconsiderados como conjunto original apenas 39 arquivos, compostos de derivadas em formatos TIFF ou JPEG e com conteúdo idêntico aos originais em formato CR2.

Na organização e descrição do acervo, optou-se por utilizar a categoria "coleção" como nível mais amplo (nível 1), tendo em vista tratar-se de um conjunto de documentos com características comuns e reunidos intencionalmente, resultado de uma missão fotográfica e exclusivamente composto de fotografias nato digitais, sem a presença de outros documentos que pudessem criar um contexto orgânico de produção documental, característico dos arquivos.

Uma vez estabelecida a natureza do conjunto documental, estabeleceu-se seu arranjo em 43 dossiês intitulados com o nome do médico retratado, que reúnem os respectivos itens organizados e descritos por ordem cronológica. Tal opção procura respeitar a lógica de produção expressa tanto nos próprios documentos quanto na publicação da qual foram objetos.

Os itens documentais são descritos por meio de um código de referência específico a cada um deles; um título que inclui pessoas, situações e locais retratados; nome do produtor; data e local de produção (informação da cidade usada como indexador); e informações sobre dimensão e suporte (gênero, formato de arquivo, padrão de cor e tamanho em pixels). A descrição da coleção, em seus diferentes níveis hierárquicos (coleção, dossiê, itens), está disponibilizada para acesso on-line na Base Arch⁸, acompanhada das imagens derivadas de acesso, em formato imagem JPEG.

Os documentos digitais submetidos à Base Arch são descritos de acordo com as seguintes normas: Norma Brasileira de Descrição Arquivística (Nobrade), Norma Internacional de Registro de Autoridade Arquivística para Entidades Coletivas, Pessoas e Famílias (ISAAR) e Norma Internacional para Descrição de Instituições com Acervo Arquivístico (ISDIAH).⁹

3.2. Aspectos organizacionais

Guarda, tratamento, preservação e acesso do acervo arquivístico permanente sob responsabilidade da COC cabem ao Departamento de Arquivo e Documentação, por meio de dois de seus serviços: o Serviço de Arquivo Histórico (SAH) e o Serviço de Conservação e Restauração de Documentos (SCRD).

O primeiro incumbe-se de receber os arquivos que ingressam no acervo, seja na forma de recolhimento, para os documentos

8. Disponível em: <http://basearch.coc.fiocruz.br/index.php/mais-medicos>; acesso em 22 mar. 2022.

9. As normas encontram-se disponíveis, respectivamente, em: http://conarq.gov.br/images/publicacoes_textos/nobrade.pdf; http://conarq.gov.br/images/publicacoes_textos/isaar_cpf.pdf; http://conarq.gov.br/images/publicacoes_textos/isdiah.pdf; acesso em 23 mar. 2022.

institucionais da Fiocruz, ou nas demais modalidades de incorporação de acervos arquivísticos históricos.

Após a entrada do arquivo e de acordo com planejamento do Departamento, ele é submetido, no SCRD, a diagnóstico do estado de conservação, higienização e acondicionamento provisório de seus documentos. Concluídos esses processos, pode-se dar início no SAH à organização do arquivo, que inclui as etapas de identificação, arranjo e descrição, bem como a inserção de representantes digitais de seus documentos, quando disponíveis, na Base Arch. A publicação do inventário do arquivo na Base Arch é o marco de sua disponibilidade para consulta. O SAH possui uma Sala de Consulta, encarregada pelo atendimento da pesquisa aos acervos permanentes. É também por meio da Sala de Consulta que pedidos de reprodução de documentos são atendidos.

A equipe responsável pelo tratamento arquivístico da Coleção Programa Mais Médicos constou de dois documentalistas, que se dedicaram à identificação das imagens e à coleta de informações contextuais de produção dos registros junto ao fotógrafo Araquém Alcântara.

A equipe do Laboratório Fotográfico, área também integrante do SAH, participa do planejamento das ações de preservação digital dos documentos fotográficos que integram os arquivos permanentes e é responsável por sua execução. Em relação às imagens analógicas, cabem à equipe as reproduções digitais destinadas à preservação e a geração de cópias para usuários.

Quanto às fotografias nato digitais, a Coleção Programa Mais Médicos acrescentou novos procedimentos de preservação digital. Após procedimentos de backup, os fotógrafos realizaram a conferência dos arquivos digitais nos HDs originais, quando se constatou que a coleção estava distribuída

em quatro pastas contendo arquivos nos formatos CR2, TIFF e JPEG. Em seguida, uma cópia de segurança de todo o conteúdo da coleção foi gerada no servidor local do laboratório. Posteriormente foram geradas mais duas cópias de segurança do acervo, para guarda externa.

3.3. Aspectos técnicos

Os documentos nato digitais da Coleção Programa Mais Médicos estão armazenados em diferentes tipos de tecnologia: HDs externos; HD interno de um computador desktop; servidor local do tipo NAS (Network-Attached Storage); e servidores externos. Tanto os HDs quanto o servidor NAS estão localizados no Laboratório Fotográfico, ao passo que os servidores no Serviço de Tecnologia da Informação (STI) e na sala-cofre situam-se em outros prédios no campus da Fiocruz.

Conforme mencionado, a coleção é composta de arquivos digitais nos formatos CR2, TIFF e JPEG. A partir dos arquivos CR2 foram geradas derivadas em TIFF e JPEG, formatos amplamente utilizados e que podem ser reproduzidos por uma grande quantidade de softwares.

As fotografias nato digitais da coleção estão organizadas e disponíveis para consulta na Base Arch, repositório de informações baseado no sistema aberto AtoM - Access to Memory¹⁰, desenvolvido pelo Conselho Internacional de Arquivos (CIA) e customizado pela COC de acordo com normas nacionais e internacionais de descrição arquivística. Funcionando em ambiente web (versão 2.3.1), o AtoM é uma aplicação de código aberto, destinado à descrição normalizada de arquivos. A base

10. Disponível em: <https://www.accesstomemory.org/pt-br/>; acesso em 25 jun. 2020.

permite também exportar metadados em formato XML para os padrões Dublin Core e EAD (Encoded Archival Description).

3.4. Requisitos dos usuários

Consideram-se usuários internos do acervo arquivístico permanente os pesquisadores e técnicos da COC, alunos dos cursos de pós-graduação e dos cursos livres ofertados, colaboradores das áreas de gestão da unidade, estagiários e bolsistas, além de profissionais da Assessoria de Comunicação (Ascom) da COC. Também são considerados usuários internos todos os colaboradores das outras unidades da Fiocruz, suas respectivas assessorias de comunicação e toda a comunidade de alunos e bolsistas da instituição. Os usuários externos são todos os que possam se beneficiar do acesso ao arquivo da COC – pesquisadores, estudantes, bem como o público em geral.

Todas as imagens da Coleção Programa Mais Médicos estão disponíveis na Base Arch, o que dispensa mediação para seu acesso. É possível navegar pelos dossiês, visualizar as fotografias e fazer seu download em baixa resolução. Cópias em alta resolução devem ser solicitadas à Sala de Consulta, por meio do serviço Consultar Arquivos Históricos da Fiocruz, no Portal Gov. Br.¹¹ De acordo com a utilização pretendida (web, livro, artigo, poster etc.), são gerados arquivos digitais seguindo parâmetros específicos, nos formatos TIFF ou JPEG, tanto em baixa como em alta resolução.

11. Disponível em: <https://www.gov.br/pt-br/servicos/consultar-arquivos-historicos-da-fiocruz>; acesso em 22 mar. 2022.

3.5. Práticas correntes de preservação

Colaboradores da COC são membros do Grupo de Trabalho em Preservação Digital do PRESERVO e do Grupo de Trabalho Gestão e Preservação de Documentos Arquivísticos Digitais, ambos dirigidos ao estabelecimento de ações e diretrizes para preservação digital de acervos. Produzido por esse último grupo e publicado pelo Sigda/Fiocruz, o documento *Padrão de metadados de documentos arquivísticos digitais da Fundação Oswaldo Cruz: manual de aplicação para a fase de produção de documentos*,¹² estabelece metadados mínimos necessários para a descrição dos documentos iconográficos digitais compostos de 15 campos, que incluem metadados descritivos e técnicos.

Em relação à Coleção Programa Mais Médicos, foram registrados os metadados dos arquivos digitais originais (formato XMP), armazenados nos HDs fornecidos por Araquém Alcântara, que contemplam informações importantes, como o modelo de câmera e o software de edição utilizados pelo fotógrafo. Além disso, metadados descritivos (como nome do produtor, data e local de produção) e técnicos (como formato de arquivo, padrão de cor e tamanho em pixels) foram criados e submetidos à Base Arch, de modo que qualquer usuário da plataforma pode ter acesso a eles.

Os arquivos digitais da coleção foram armazenados em HDs instalados nas estações de trabalho do Laboratório Fotográfico e, após a etapa de organização do acervo, copiados em um servidor local do tipo NAS – que pode ser acessado por todos os fotógrafos via rede. Conforme mencionado anteriormente, os documentos foram também copiados para dois ambientes externos: o servidor local da STI e a sala-cofre da Fiocruz. Essa estratégia de backup dos documentos digitais da Coleção

12. Disponível em: http://www.sigda.fiocruz.br/images/pdf/Manual_Padrao_Metadados_SIGDA_VERSAO_FINAL_JUNHO_2020.pdf; acesso em 22 mar.2022.

Programa Mais Médicos segue a regra de preservação digital conhecida como “3-2-1”: a produção de ao menos três cópias de cada documento, mantidas em dois tipos diferentes de armazenamento (HDs e servidores), com pelo menos um deles separado dos demais.

4. AVALIAÇÃO DE RISCOS

A fim de fornecer as bases para a elaboração de uma estratégia de preservação digital, é recomendável avaliar previamente os riscos que possam impactar negativamente a gestão e a preservação da coleção Programa Mais Médicos, além de indicar ações de tratamento possíveis. Os Quadros 1 e 2, apresentados a seguir, baseiam-se em modelo e critérios apresentados no *Guia de gestão de riscos da Fiocruz*¹³; por sua vez, as áreas funcionais indicadas no Quadro 2 baseiam-se no *Levels of Digital Preservation*¹⁴ (Níveis de Preservação Digital) – conjunto hierárquico de recomendações desenvolvido pela Library of Congress.

A avaliação preliminar apresentada no Quadro 2 indica níveis de risco de médio a alto, associados à preservação dos documentos digitais da coleção em questão. Os níveis mais acentuados de risco, que obrigatoriamente requerem ação mais imediata, envolvem três fatores: 1) perda ou corrupção dos arquivos digitais; 2) alteração inesperada ou deliberada dos arquivos digitais; e 3) acesso não autorizado aos arquivos digitais. Por sua vez, os riscos de nível médio giram em torno de dois fatores:

13. Disponível em: https://www.arca.fiocruz.br/bitstream/icict/39359/2/guia_de_gestao_de_riscos_da_fiocruz.pdf; acesso em 23 mar. 2022.


14. Disponível em: <https://ndsa.org/publications/levels-of-digital-preservation/>. acesso em: 23 mar. 2022.

1) metadados insuficientes; e 2) obsolescência de formatos de arquivo e dispositivos de armazenamento. No mesmo quadro são apontados também os possíveis impactos associados a cada um dos riscos, os controles existentes para reduzir o seu impacto, além das possíveis ações de tratamento para a sua mitigação.

Quadro 1 – Níveis de Risco

IMPACTO	Muito Alto 10	10 RM	20 RM	50 RA	80 RE	100 RE
	Alto 8	8 RB	16 RM	40 RA	64 RA	80 RE
	Médio 5	5 RB	10 RM	25 RM	40 RA	50 RA
	Baixo 2	2 RB	4 RB	10 RM	16 RM	20 RM
	Muito Baixo 1	1 RB	2 RB	5 RB	8 RB	10 RM
		Muito Baixa 1	Baixa 2	Média 5	Alta 8	Muito Alta 10
	PROBABILIDADE					

 RB: Risco Baixo

 RM: Risco Médio

 RA: Risco Alto

 RE: Risco Muito Alto/Estratégico

Quadro 2 – Planilha de identificação e avaliação de riscos

Área Funcional	Risco	Consequências/ Impactos	Nível de Risco	Controles Existentes	Ações de Tratamento
Armazenamento	Perda ou corrupção dos arquivos digitais	Incapacidade de ler, acessar e recuperar conteúdos.	ALTO Probabilidade (5) X Impacto (8) = 40	Os arquivos digitais da coleção são mantidos em diferentes tipos de armazenamento (HDs externos, servidor e fitas LTO), havendo redundância e dispersão geográfica (cópia de segurança localizada na sala-cofre do campus de Manginhos).	Documentar as mídias de armazenamento utilizadas, a fim de assegurar sua longevidade, indicando os recursos e requisitos necessários para o seu funcionamento.
Integridade	Alteração inesperada ou deliberada dos arquivos digitais	Perda de fixidez dos arquivos digitais, comprometendo a sua integridade. A capacidade de manutenção da integridade é um dos elementos essenciais da preservação digital.	ALTO Probabilidade (8) X Impacto (8) = 64	Só há verificações de fixidez e de presença de vírus nos arquivos digitais que estão armazenados no servidor da STI. Não há procedimentos desse tipo nos dispositivos localizados no laboratório.	Geração de soma de verificação (checksum) de todos os arquivos digitais, além de sua verificação a cada seis meses, no mínimo. Os valores de checksum deverão ser gerados e verificados em todos os dispositivos onde estão armazenados os arquivos digitais da coleção. Além disso, realizar verificação de vírus, rotineiramente, em todos os dispositivos.
Controle	Acesso não autorizado aos arquivos digitais	Corrupção, alteração ou exclusão dos arquivos digitais.	ALTO Probabilidade (5) X Impacto (8) = 40	O acesso aos arquivos digitais armazenados nos HDs externos e no servidor local é restrito aos fotógrafos do laboratório, mas não é formalmente documentado.	Documentar todos os profissionais que terão autorização para ler, editar e mover os arquivos digitais da coleção.

Área Funcional	Risco	Consequências/ Impactos	Nível de Risco	Controles Existentes	Ações de Tratamento
Metadados	Metadados insuficientes	Impossibilidade de descrever, compreender, encontrar e recuperar o conteúdo adequadamente. Metadados consistentes são cruciais para prover presunção de autenticidade aos conteúdos, incluindo informações sobre as ações de preservação realizadas ao longo do ciclo de vida dos arquivos digitais.	MÉDIO Probabilidade (5) X Impacto (5) = 25	Alguns metadados descritivos e técnicos foram copiados dos arquivos originais e exportados para a base Arch, mas são insuficientes para descrever plenamente os conteúdos.	Geração de metadados técnicos para todos os documentos digitais da coleção.
Conteúdo	Obsolescência de formatos de arquivo e dispositivos de armazenamento	Incapacidade de reprodução e leitura dos arquivos digitais, comprometendo o acesso aos seus conteúdos.	MÉDIO Probabilidade (5) X Impacto (5) = 25	A partir dos arquivos digitais originais produzidos pela câmera de Araquém Alcântara, foram gerados arquivos TIFF – formato muito utilizado como matriz de preservação digital. Porém, não há um monitoramento de obsolescência nem procedimentos de migração previstos.	Monitoramento permanente dos processos de obsolescência e mudanças tecnológicas, além da execução de migrações de formato e outras ações que assegurem o acesso ao conteúdo.

5. PLANEJAMENTO DE PRESERVAÇÃO

5.1. Prioridades de preservação

A estratégia de preservação digital adotada na Coleção Programa Mais Médicos pode ser estendida a todos os documentos natos digitais, desde que sejam observadas as particularidades referentes aos metadados de cada gênero documental. Contudo, a prioridade de preservação deve ser dada aos documentos arquivísticos em idade permanente, já organizados e descritos arquivisticamente.

5.2. Estratégia de preservação avaliada

A estratégia de preservação avaliada envolve a criação de pacotes de informação para submissão (submission information package– SIP), associada à implementação de um conjunto de ações de preservação digital – desde a verificação de integridade dos arquivos digitais até o seu armazenamento adequado – que compõem um fluxo de trabalho estruturado.

Em síntese, essa estratégia tem três objetivos inter-relacionados:

- Conferir presunção de autenticidade aos documentos digitais;
- Estar em consonância com o modelo de referência Open Archival Information System (OAIS)¹⁵, por meio da preparação de pacotes de informação;
- Criar fluxos de trabalho sustentáveis para a incorporação e a preservação de documentos natos digitais.

15. O modelo de referência OAIS é recomendação internacional desde 2003 (ISO 14721). Trata-se de um modelo conceitual que define um repositório digital identificando o ambiente, os componentes funcionais, suas interfaces internas e externas, os objetos de dados e informações.

No contexto do modelo de referência OAIS, um pacote de informação é um recipiente conceitual composto por dois tipos de dados: os documentos digitais (informação de conteúdo) e seus metadados (informação de representação), além de uma soma de verificação (checksum)¹⁶. Um pacote SIP [Figura 1] é criado visando à futura admissão dos documentos digitais e seus metadados associados em um ambiente de preservação tal como um repositório digital.¹⁷

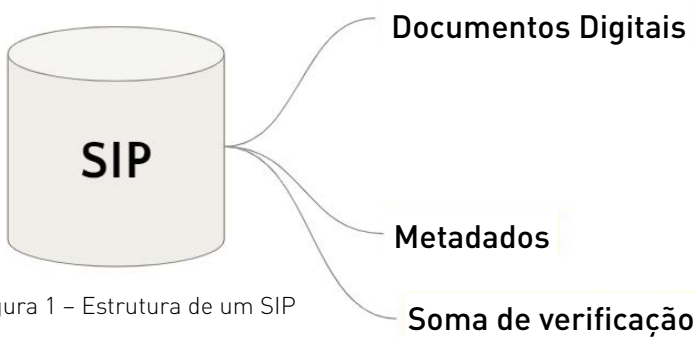


Figura 1 – Estrutura de um SIP

16. Uma soma de verificação (checksum) é uma sequência de caracteres alfanuméricos, que resulta da aplicação de um algoritmo em um arquivo digital. A soma de verificação permanecerá a mesma enquanto o arquivo se mantiver inalterado, de forma que pode ser usada para identificar arquivos alterados, corrompidos ou duplicados.

17. Além do SIP, o modelo OAIS estabelece mais dois tipos de pacotes de informação: o pacote de informação para arquivamento (archival information package – AIP), que se refere ao acondicionamento e armazenamento dos documentos digitais e seus metadados associados; e o pacote de informação para disseminação (dissemination information package – DIP), que se refere ao acesso aos documentos digitais e seus metadados. Apenas o SIP está dentro do escopo deste documento, que se limita a traçar uma estratégia de preservação digital que independe da adoção de um repositório digital.

Através da criação de SIPs, este plano de preservação pretende traçar uma estratégia que não dependa da adoção de um repositório digital ou de sistema específico de preservação digital.¹⁸ Mesmo que um SIP ainda não esteja inserido em um repositório digital, sua estrutura já fornece as bases para a preservação digital dos documentos, ao conferir atributos que asseguram integridade e presunção de autenticidade aos documentos – incluindo um conjunto de metadados e uma soma de verificação dos arquivos.

Além da criação de pacotes de informações, outras ações também deverão ser colocadas em prática a fim de contemplar requisitos essenciais de preservação digital, como, por exemplo, a extração de metadados técnicos embutidos nos arquivos digitais e a criação de somas de verificação.

5.3. Funções e responsabilidades

Para a implementação do plano, estarão envolvidos pelo menos três profissionais do DAD, responsáveis pelas ações de incorporação e preservação dos documentos nato digitais. Caberá ao Serviço de Tecnologia da Informação (STI) da COC a responsabilidade de executar periodicamente as cópias de segurança dos documentos. Já a Coordenação Geral de Gestão de Tecnologia de Informação (Cogetic) da Fiocruz deverá ficar responsável pela replicação periódica dos dados digitais armazenados no servidor local da COC para servidor próprio, além de se incumbir da gestão e preservação dos mesmos.

18. Uma vez que algumas unidades da Fiocruz vêm realizando testes com o Archivemática, sistema de gestão de preservação digital, este software poderá vir a ser utilizado para a admissão dos SIPs.

5.4. Custos e fontes de financiamento

Os custos necessários para a implementação do plano de preservação digital deverão ser levantados conjuntamente pela COC – por meio do STI – e a Cogetic. Recomenda-se que contemplem tanto infraestrutura tecnológica quanto recursos humanos, além da previsão de investimentos necessários na atualização contínua dos registros preservados (incluindo hardwares, softwares e sistemas operacionais), de modo a dirimir os processos de obsolescência tecnológica.

As principais fontes de financiamento para o plano devem incluir, sobretudo, recursos advindos do orçamento da Fiocruz, além de editais de fomento e leis de incentivo dedicados ao patrimônio

6. AÇÕES DE PRESERVAÇÃO

Segue a descrição das etapas relativas à preservação dos documentos nato digitais, com os seus respectivos fluxos de trabalho incluindo, de maneira sequencial, as ações específicas necessárias para a sua realização. Para o cumprimento de determinadas ações, serão relacionadas ferramentas abertas e gratuitas¹⁹ que poderão ser utilizadas para facilitar sua execução.

- **Estruturar os arquivos digitais em pacote de informação.**
Estruturar e criar pacote de informação contendo documentos digitais e metadados. O pacote deverá conter

19. Ferramentas abertas ou de código aberto são aquelas em que o código-fonte está disponível e publicamente documentado, além de haver uma comunidade de usuários e desenvolvedores que as apoia. Embora as ferramentas abertas sejam geralmente de uso gratuito, isso nem sempre significa que elas não requeiram outros recursos para sua implementação e uso.

duas pastas: “objects”, que abriga os documentos digitais a serem preservados; e “metadata”, que abriga os metadados coletados. A pasta “metadata” deverá conter a subpasta “submissionDocumentation”, que abriga documentos relacionados.

- **Copiar metadados técnicos embutidos nos arquivos digitais.**
Os metadados técnicos embutidos nos arquivos digitais deverão ser copiados e exportados para um arquivo em formato XML (a fim de facilitar futura admissão em um sistema de preservação digital). O arquivo deverá ser incluído na subpasta “submissionDocumentation” do pacote de informação.
Ferramenta gratuita: [ExifTool](#).
- **Criar planilha de metadados técnicos e descritivos.**
Gerar metadados técnicos e descritivos de acordo com as recomendações apresentadas no manual *Padrão de metadados de documentos arquivísticos digitais da Fundação Oswaldo Cruz*, mencionado anteriormente. Os metadados deverão ser preenchidos em um documento de planilha, formato CSV.
Ferramenta gratuita: [LibreOffice Calc](#).
- **Gerar soma de verificação dos arquivos digitais.** Gerar soma de verificação (checksum) para cada arquivo digital, a fim de garantir a sua integridade. O arquivo contendo o conjunto das somas de verificação (formato MD5) deverá ser incluído na pasta “metadata”. Validar os pacotes ao menos uma vez ao ano, a fim de verificar se os arquivos sofreram alguma alteração, corrompendo sua integridade.
Ferramenta gratuita: [MD5Checker](#).

- **Gerar e documentar cópias de segurança (backup).** Gerar, no mínimo, duas cópias de segurança: uma que esteja localizada próxima à área de trabalho, a ser dedicada à recuperação rápida de arquivos perdidos ou corrompidos; e outra geograficamente separada das demais, para o caso de perda substancial de arquivos. Deve-se utilizar uma ferramenta confiável para a cópia dos arquivos. Essa estratégia de geração de cópias deverá ser documentada e constantemente atualizada.
Ferramenta gratuita: [Robocopy](#).
- **Documentar todos os profissionais que terão autorização para ler, editar e mover os arquivos digitais.** Atribuir, registrar e gerenciar permissões, a fim de atenuar o risco de corrupção ou alteração de dados. Elaborar também uma trilha de auditoria, com registro de todas as ações que cada usuário realizou com os arquivos digitais. Enquanto não houver um sistema automatizado para gerenciamento das permissões e para registro da trilha de auditoria, essas ações deverão ser executadas manualmente, baseadas no modelo apresentado no Apêndice B – que deverá estar disponível para todos os profissionais habilitados, via intranet.

7. REVISÃO DO PLANO DE PRESERVAÇÃO

Este documento deverá ser revisto em período não superior a dois anos a partir da data de sua última versão, podendo ser atualizado de acordo com a necessidade. Alguns eventos, contudo, poderão exigir a sua revisão antecipada, tais como:

- Alterações no orçamento previsto para a preservação digital, podendo gerar mudanças de investimentos e custos relacionados;
- Adoção de um sistema de gestão e preservação digital;
- Alterações na tecnologia adotada em uma ou mais etapas do fluxo de trabalho de preservação digital;
- Mudanças na estrutura organizacional ou no número de profissionais envolvidos com a preservação digital.

APÊNDICE A – INSTRUÇÕES DE USO DAS FERRAMENTAS GRATUITAS

ExifTool

O ExifTool é um software de código aberto que utiliza linha de comando e é voltado para leitura, importação e exportação de metadados de imagem, áudio e vídeo.

Para instalar o ExifTool:

1. Faça o download do Windows Executable (arquivo “exiftool-12.01.zip”) na página do ExifTool: <https://exiftool.org/index.html>.
2. Extraia “exiftool (-k) .exe” do arquivo .zip e transfira-o para a área de trabalho.
3. Renomeie o arquivo “exiftool (-k) .exe” para “exiftool.exe”.
4. Transfira o arquivo “exiftool.exe” para o diretório “C:\WINDOWS”.

Para executar o ExifTool:

1. Abra o prompt de comando do DOS clicando com o botão direito do mouse no botão “Iniciar” do Windows e selecionando a opção “Executar”; na caixa de diálogo “Abrir”, digite “cmd” e dê ok.
2. Para copiar os metadados de todos os arquivos de um diretório, incluindo subdiretórios, e gerar um arquivo XML, digite um comando seguindo o modelo abaixo:

Exemplo: exiftool “c:\DAD\Mais Medicos” -s -G -r -csv →metadata_exif.csv

exiftool [Unidade]:\[Caminho_do_diretório] -s -G -r -X → [Nome_arquivo_saída.xml]

Exemplo: exiftool “c:\DAD\Mais Medicos” -s -G -r -X →metadata_exif.xml

3. Feche o prompt de comando do DOS.
4. Salve o arquivos gerado “metadata_exif.xml”) na subpasta “submissionDocumentation”.

Descrição dos comandos utilizados:

“[Unidade]:\[Caminho_do_diretório]” é a unidade e o caminho onde o diretório está localizado. Os caminhos que contiverem espaços deverão estar entre aspas.

“-s” exibe um padrão de saída mais resumido.

“-G” exibe o nome do grupo para cada elemento de metadados.

“-r” processa todos os subdiretórios.

“→” indica que será gerado um arquivo de saída.

“-X” exporta os elementos em formato XML.

“[Nome_arquivo_saída.xml]” é o nome que deve ser atribuído ao arquivo a ser gerado (formatos XML).

Instruções adicionais sobre o uso do ExifTool: <https://www.weareavp.com/exiftool-tutorial-series/>.

MD5Checker

O MD5Checker é um software gratuito para gerar e validar somas de verificação (checksums) de arquivos digitais. O download está disponível na página do desenvolvedor: <http://getmd5checker.com>.

Para gerar somas de verificação:

1. Abra o programa e clique no botão “Add”.
2. Na nova janela que se abre, navegue até o diretório desejado, selecione todos os arquivos e clique no botão “Abrir”.
3. De volta à janela original, clique no botão “Save” para salvar o checksum de cada arquivo.

4. Clique no botão “S To”. Na nova janela que se abre, marque a opção “Two spaces”. Nomeie o arquivo como “checksum.md5”, selecione o diretório de destino (“metadata”) e clique em “Salvar”.

Para validar somas de verificação:

1. Abra o programa e clique no botão “Add”.
2. Na nova janela que abre, selecione o arquivo de soma de verificação (“checksum.md5”) a ser validado e clique em “Abrir”.
3. Caso não sejam detectados erros, um ícone verde aparecerá ao lado de cada um dos arquivos digitais. Caso erros sejam detectados, um ícone vermelho aparecerá ao lado de cada um dos arquivos digitais; nesse caso, será necessário substituí-los pelas suas cópias de segurança.

Robocopy

O Robocopy é um comando do Windows que permite copiar rapidamente grandes quantidades de arquivos digitais, mantendo todos os seus atributos. O Robocopy está disponível em todas as versões do Windows e pode ser executado no prompt de comando do DOS. Não é necessário instalar nenhum software adicional.

1. Para abrir o prompt de comando do DOS, clique com o botão direito do mouse no botão “Iniciar” do Windows e selecione a opção “Executar”; na caixa de diálogo “Abrir”, digite “cmd” e dê ok.

Para copiar os arquivos digitais, digite um comando seguindo o modelo abaixo:

```
robocopy [Unidade_de_origem]:\[Diretório_de_origem] [Unidade_de_destino]:\  
[Diretório_de_destino] /MIR /DCOPY:T
```

Exemplo: robocopy “c:\Mais Medicos” e:\DAD /MIR /DCOPY:T

2. Quando a cópia estiver concluída, digite “exit” e pressione a tecla Enter para fechar a janela do DOS.

Observação: Sempre copie os arquivos para um diretório de destino que ainda não exista, pois o comando “/MIR” é capaz de excluir arquivos em um diretório de origem quando um caminho de destino preexistente é usado.

Descrição dos comandos utilizados:

“[Unidade_de_origem]:\[Diretório_de_origem]” é a unidade (drive) e o caminho onde os arquivos originais estão localizados. Os caminhos que contiverem espaços deverão estar entre aspas.

“[Unidade_de_destino]:\[Diretório_de_destino]” é a unidade e o caminho para onde os arquivos deverão ser copiados. Certifique-se que esse caminho ainda não exista.

“/MIR” faz com que o Robocopy copie todos os arquivos e subdiretórios exatamente da forma como aparecem.

“/DCOPY:T” faz com que o Robocopy preserve os timestamps (informações de hora e de data) em todos os arquivos, diretórios e subdiretórios.

Instruções adicionais sobre o uso do Robocopy: <https://docs.microsoft.com/pt-br/windows-server/administration/windows-commands/robocopy>.

APÊNDICE B – PERMISSÃO DE ACESSO E TRILHA DE AUDITORIA

As tabelas abaixo são modelos ilustrativos e podem ser ajustadas a fim de melhor refletir as atribuições e responsabilidades dos usuários.

Usuários	Permissão de acesso			
	Ler	Editar	Mover	Apagar
Profissionais do DAD	X	X	X	X
Profissionais do STI	X	X	X	
Outros usuários	X			

Usuário	Trilha de auditoria		
	Arquivo	Ação	Data
José da Silva	IMG_0047.tiff	Moveu da unidade C para a unidade E	2020-03-23



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

