

Preservação de documentos audiovisuais avaliação da aplicabilidade do Archivematica

João Guilherme Nogueira Machado ¹Miguel Angel Mardero Arellano ²Cleomar Huche Lopes ³

Resumo

O entendimento teórico sobre o que os repositórios digitais se propõem a preservar vem do aprendizado prático de uso e interpretação das ações geradas por esses *softwares* cada vez mais indispensáveis. A preservação de documentos audiovisuais e suas especificidades, sobretudo a variedade de formatos físicos, formatos digitais de contêineres e *codecs*, assim como seu alto volume de dados faz com que esse gênero documental mereça uma atenção diferenciada no seu tratamento arquivístico de preservação. O trabalho aqui proposto utilizou como ferramenta de preservação a ser testada o repositório Archivematica na versão 1.9.1, no contexto do seu uso prático para tratamento de documentos audiovisuais que compõem o acervo da VideoSaúde – Distribuidora (VSD) da Fiocruz. Os resultados obtidos foram descritos e avaliados quanto a sua capacidade de criação de pacotes de informação de submissão (SIP), geração de pacotes de informação de arquivamento (AIP) e interpretação de metadados METS dos AIP. Nesse sentido, concluiu-se que a estrutura dos SIP elaborados para teste no *software* atendeu aos parâmetros de processamento para geração de AIP completos e robustos. Entretanto, destaca-se que testes com outros formatos e abordando outras especificidades de vídeo devem ser pesquisados para complementação da análise sobre o uso da ferramenta frente à preservação digital de documentos audiovisuais.

Palavras-chave: Audiovisual. Preservação digital. Repositórios. Metadados.

¹ Fundação Oswaldo Cruz – Rio de Janeiro, RJ / Brasil

E-mail: guilherme.machado@icict.fiocruz.br

² Instituto Brasileiro de Informação em Ciência e Tecnologia – Brasília, DF / Brasil – E-mail: miguel@unicamp.br

³ Fundação Oswaldo Cruz – Rio de Janeiro, RJ / Brasil

E-mail: cleomar.lopes@icict.fiocruz.br

Submetido em: 10/09/2020

Aceito em: 13/11/2020

Publicado em: 18/11/2020

Preservation of audiovisual files evaluation of the aplicability of Archivematica

Abstract

The use of repositories to preserve digital documents is based on the recommendations of the Open Archival Information System (OAIS) reference model and is consolidated as a fundamental part of technological solutions for maintaining the authenticity of digital archival documents. The theoretical understanding of what these preservation repositories propose to accomplish comes from the practical learning of using and interpreting the actions generated by these increasingly indispensable software. The preservation of audiovisual documents and their specificities, especially the variety of physical formats, digital formats of containers and codecs, as well as its high volume of data makes this documentary genre merit a differentiated attention in its archival preservation treatment. The work proposed here used the Archivematica repository in version 1.9.1 as a preservation tool to be tested, in the context of its practical use for processing audiovisual documents that make up the collection of VideoSaúde – Distribuidora (VSD) of Fiocruz. Digital preservation literature and software documentation were consulted to support the development of the tests. The results obtained were described and evaluated for their ability to create submission information packages (SIP), generation of archival information packages (AIP) and interpretation of AIP METS metadata. In this sense, it was concluded that the structure of the SIP designed for testing in the software met the processing parameters for generating complete and robust AIP. However, it is highlighted that tests with other formats and addressing other specificities of video should be researched and / or carried out to complement the analysis on the use of the tool in view of the digital preservation of audiovisual documents.

Keywords: Audiovisual. Digital preservation. Repositories. Metadata.



1 Introdução

As inovações na área das Tecnologias Digitais da Informação e Comunicação – TDCIs são capazes de traduzir os contínuos avanços da informática e sua fundamental participação na construção de conhecimentos e tecnologias voltadas para a preservação e acesso ao patrimônio científico e cultural digital. Nesse contexto, a formação de acervos híbridos compostos por documentos analógicos e digitais está cada vez mais ligada à realidade de diversas instituições sejam elas públicas ou privadas.

Segundo Cavalcante (2007, p.154):

A digitalização desses documentos e a criação de um patrimônio digital, além do fascínio que exerce e da justificação social, histórica e cultural que apresenta necessita, entretanto, de um importante trabalho de reflexão, que possa ir além de leitura linear ocasionada pela euforia do presente. Reflexão que inicia com a constatação de tratar-se de uma representação do real, especialmente quando nos referimos a documentos históricos, (representativos de fenômenos sociais ou testemunhas de uma era), produzidos no passado, em suportes físicos e atualmente digitalizados, que não poderão efetivamente (pelo menos no presente) substituir a obra em si, em sua materialidade, mesmo que se configurem em reproduções de alta qualidade e de perfeição completamente idênticas aos originais, de uma fidelidade impressionante. (CAVALCANTE, 2007, p.154)

Não obstante a essa condição, que recai com maior fundamentação aos documentos arquivísticos, encontra-se o gênero documental objeto do estudo: o audiovisual.

Por documentos audiovisuais podemos entender: gravações visuais (com ou sem banda de som [soundtrack]) independente [da sua base física] do seu suporte e processo de gravação usado, como filmes, [filmstrips] diafilme, microfilmes, diapositivos, fitas magnéticas, cinescópios [kinescopes], videogramas [videograms], videotapes - fitas de vídeo (videotape, videodiscos), discos ópticos legíveis por laser (a) planejados para recepção pública quer através de televisão ou por meio de projeção em écrans ou por quaisquer outros meios (b) destinados a ser postos à disposição do público (EDMONDSON, 2017, p.26).

Ainda segundo o pesquisador Edmondson:

o reconhecimento dos arquivos audiovisuais não se deu de maneira imediata e formal. Pelo caráter difuso das instituições mantenedoras desses tipos de documentos e pela extensão de sua utilização, o audiovisual, nesses casos, viu o crescimento de sua popularidade se desenvolver com certo atraso. (EDMONDSON, 2017, p.19)

Com frequência as instituições custodiadoras desse gênero documental se veem frente às questões ligadas a preservação e acesso ao conteúdo registrado nesses documentos, tendo que enfrentar riscos e altos custos associados ao acesso; nesse sentido, objetivar a preservação sem garantir o acesso aos materiais preservados pode ser contraditório.

Esse gênero documental também é afetado pelo impacto das constantes mudanças tecnológicas e dos efeitos mercadológicos da indústria. Para a existência e manutenção de acervos audiovisuais ações de preservação são fundamentais devido às especificidades e variedades de suportes, formatos e meios de armazenamento de conteúdos analógico e digital e suas possíveis instanciações¹. Algumas estratégias de preservação e acesso a longo prazo de documentos audiovisuais como: a conservação preventiva dos suportes físicos que armazenam os conteúdos, a digitalização desses documentos com o suporte de repositórios digitais de preservação e robustos perfis de aplicação de metadados tem destaque na literatura mais recente.

Diversas possibilidades e estratégias de preservação digital estão em constante desenvolvimento e, dentro desse quadro, torna-se cada vez mais evidente a necessidade de colocá-las em prática e testar suas potencialidades e fragilidades. Impõe-se como fundamental o levantamento das mais recentes e consolidadas teorias e práticas acerca da preservação digital que proporcionem, especialmente a possíveis produtores de conteúdo audiovisual, a elaboração e construção de pacotes de informação para submissão em repositórios digitais de preservação de acordo com recomendações do modelo conceitual de referência *Open Archival Information System* – OAIS (ISO, 2012). O modelo de referência OAIS tornou-se em 2003 uma norma internacional, ISO 14721: 2003, com o

objetivo de estabelecer um sistema de arquivamento de informações, por meio de um esquema organizacional composto de pessoas que aceitam a responsabilidade de preservar as informações e disponibilizá-las para uma comunidade designada. Vale ressaltar que as recomendações implícitas no OAIS dão ênfase à presunção de autenticidade, integridade e interoperabilidade dos pacotes. (CCSDS, 2012).

No Brasil, o Conselho Nacional de Arquivos (CONARQ) publicou em setembro de 2015 a sua resolução de nº 43 que dispõe sobre a adoção de repositório arquivístico digital confiável (RDC-Arq), em consonância com o modelo de referência OAIS, enquanto sistema para “o arquivamento e manutenção dos documentos arquivísticos em suas fases corrente, intermediária e permanente em formato digital, e de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade e a preservação desses documentos”

2 Justificativa

De acordo com o documento traduzido referente ao Projeto Interpares com o título "Diretrizes do Preservador - A Preservação de Documentos Arquivísticos digitais: Diretrizes para Organizações" (Arquivo Nacional/Câmara dos Deputados, 2010), aspectos relacionados à manutenção da autenticidade e integridade dos documentos

¹ Uma manifestação de um ativo incorporado em formato físico ou digital, como uma fita, DVD ou arquivo digital. Um ativo pode ter muitas instanciações, mas geralmente, cada instanciação possui o mesmo conteúdo intelectual.

digitais, segundo alguns estudos sobre da área de preservação digital, tornam-se fundamentais para os processos de produção, tramitação e arquivamento permanente de documentos arquivísticos digitais. Levando em conta esses aspectos e objetivando seguir as recomendações das resoluções nº 37 e nº 43 do Conselho Nacional de Arquivos (CONARQ, 2012 e 2015) e tendo como ponto de referência o modelo conceitual OAIS – *Open Archival Information System*, o GT de Gestão e Preservação de Documentos Arquivísticos Digitais em ação conjunta e coordenada pelo PRESERVO – Complexo de Acervos da Fiocruz, decidiu em 2018 pela implantação do Archivematica enquanto ferramenta de preservação digital dos seus acervos. A fundamentação prática da escolha foi baseada nas informações coletadas² sobre o funcionamento do *software* que fornece um conjunto integrado de ferramentas gratuitas e de código aberto que permitem aos usuários processar objetos digitais, desde a ingestão até o armazenamento e acesso ao arquivo, em conformidade com o modelo funcional [OAIS](#) e outros padrões e práticas recomendadas de preservação digital.

O trabalho aqui proposto, tem como base empírica o acervo de vídeos da VSD do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict) da Fiocruz. A VideoSaúde foi designada para realizar os testes iniciais com o Archivematica exclusivamente com conteúdo audiovisual digital.

A VSD custodia um significativo acervo audiovisual arquivístico de caráter permanente, segundo o código³ de classificação e a tabela⁴ de temporalidade e destinação de documentos para atividades finalísticas da Fiocruz, e também uma vasta coleção de títulos especializados em saúde que retratam a construção da história da saúde pública no Brasil. O acervo se constitui de uma grande quantidade de material em meio analógico (fitas analógicas em diversos formatos) e vem se expandindo com altos volumes de material nato digital ano após ano. Sendo um acervo de características híbridas, tanto no que se refere à natureza dos materiais que custodia (arquivo e coleções temáticas) quanto aos suportes e meios (fitas magnéticas, discos ópticos, meios analógico e digital) que registram os seus conteúdos, as especificidades para se lidar com essas características confirmam a importância desse estudo sob a ótica da preservação digital e acesso por longo prazo aos materiais deste acervo institucional.

O objetivo deste trabalho é apresentar uma descrição dos testes realizados (e seus resultados) com pacotes de informação elaborados com arquivos de vídeo e seus metadados descritivos e técnicos gerados na criação de pacotes de informação de submissão (SIP) para admissão no Archivematica/ Fiocruz. Os testes, objeto do

² As informações encontram-se no próprio site do Archivematica. Disponível em:

<https://www.archivematica.org/pt-br/>

³ http://www.sigda.fiocruz.br/images/pdf/Codigo_de_classificacao_de_documentos_de_arquivo_da_Fiocruz_MEIO_FIM_dez2018.pdf

⁴ http://www.sigda.fiocruz.br/images/pdf/TTD_PARA_ATIVIDADES_FIM_DA_FUNDACAO_OSWALDO_CRUZ.pdf

trabalho, se deram no âmbito da VSD⁵ no período compreendido entre 09/10/2019 a 09/12/2019.

3 Contextualização do cenário institucional

A Fiocruz é uma instituição brasileira de pesquisa em saúde pública e é vinculada ao Ministério da Saúde. A instituição possui uma grande variedade de acervos científicos e culturais produzidos em função da sua vasta gama de atividades em pesquisa na área da saúde. Dentre outras competências da instituição vale destacar, para o objetivo do trabalho aqui proposto, “preservar, valorizar e divulgar o patrimônio histórico, cultural e científico da FIOCRUZ e contribuir para a preservação da memória das áreas de saúde e de ciências biomédicas.” (FIOCRUZ, 2016).

Nesse contexto, a instituição se apresenta como estratégica para o estado brasileiro e de excelência na área de pesquisa, atuando como agente fundamental na criação, execução e inovação de políticas públicas em saúde.

Entendidas e tratadas pela instituição como um direito humano fundamental, a comunicação e a informação científica e tecnológica tornam-se objeto de estudo e demandam pelo desenvolvimento de políticas e estratégias de preservação capazes de promover a manutenção de acesso aos conteúdos informacionais gerados pela fundação. Nesse sentido, o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde da Fiocruz (Icict) destaca ao audiovisual um papel central na comunicação e disseminação da informação em saúde, além de potencial construído da memória da saúde pública brasileira.

Políticas institucionalizadas como a de Acesso Aberto ao Conhecimento (FIOCRUZ, 2014) e a de Preservação dos Acervos Científicos e Culturais (FIOCRUZ, 2018), se comprometem a dar o suporte necessário para que a Fiocruz desenvolva estudos e estratégias de preservação com o objetivo de consolidar diretrizes e recomendações contidas nas suas políticas, como a difusão do conhecimento produzido pela instituição, a preservação e o acesso (aberto) por longo prazo ao seu patrimônio audiovisual em saúde.

A intensificação das discussões acerca das estratégias de preservação do patrimônio audiovisual da Fiocruz no âmbito do PRESERVO – Complexo de Acervos da Fiocruz, instância responsável por dirimir as políticas, programas e ações de preservação dos acervos científicos e culturais da instituição, e as ações desenvolvidas pela VSD, relatam novas perspectivas práticas sobre o uso de repositórios de preservação para documentos arquivísticos audiovisuais.

⁵ A VídeoSaúde Distribuidora (VSD) é um serviço vinculado ao Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict/Fiocruz). Dentre as suas atribuições, destaca-se a veiculação de conteúdo audiovisual no campo da saúde coletiva, em emissoras públicas, educativas e comunitárias. Atualmente com exibição na TV UFES, TV Floripa, UFPR TV, TV Câmara, TVE Bahia e Canal Saúde.

Os usos cada vez mais demandados ao audiovisual pela Fiocruz podem ser vistos e refletidos através dos diversos canais de comunicação e acesso à informação que utilizam vídeos para divulgar o conhecimento produzido pela instituição, assim como para registrar a memória da fundação (<https://portal.fiocruz.br/producao-audiovisual>). É, portanto, inegável o poder informativo e de registro histórico e cultural contido intrinsecamente na linguagem audiovisual.

Desta maneira, evidenciou-se a necessidade de a instituição realizar estudos mais aprofundados sobre estratégias de preservação digital do patrimônio científico e cultural por ela gerado. No âmbito dos estudos realizados, as discussões e proposições geradas pelo GT de Gestão e Preservação de Documentos Arquivísticos Digitais⁶ da fundação seguem as orientações e recomendações estabelecidas pelo CONARQ. A VSD é integrante do GT e representa o audiovisual enquanto gênero documental arquivístico constituinte do patrimônio científico e cultural da Fiocruz.

A VSD é um serviço vinculado ao Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict/Fiocruz). Dentre as suas atribuições, destaca-se a veiculação de conteúdo audiovisual no campo da saúde coletiva, em emissoras públicas, educativas e comunitárias (<https://portal.fiocruz.br/videosaude-distribuidora>). A VSD conta com 21 colaboradores, entre: servidores, estagiários e bolsistas.

A Distribuidora é um espaço de guarda, produção, captação, preservação, pesquisa, fomento, distribuição e circulação de material audiovisual em saúde produzido pela Fiocruz, por outras instituições e por produtores independentes, com o objetivo de contribuir para o fortalecimento do Sistema Único de Saúde (SUS) e para a melhoria das condições de vida e saúde da população brasileira. Os colaboradores da VSD apoiam atividades ligadas às três grandes áreas da distribuidora, que são: Acervo, Produção e Distribuição.

4 Metodologia

O presente estudo de caso, de caráter exploratório, apresenta de forma descritiva os aspectos técnicos da aplicação de testes de admissão de conteúdo arquivístico audiovisual (nato digital) da VSD da Fiocruz na ferramenta de preservação digital Archivematica. A metodologia aplicada visa relatar e expor de forma detalhada, aspectos técnicos e a prática de uso.

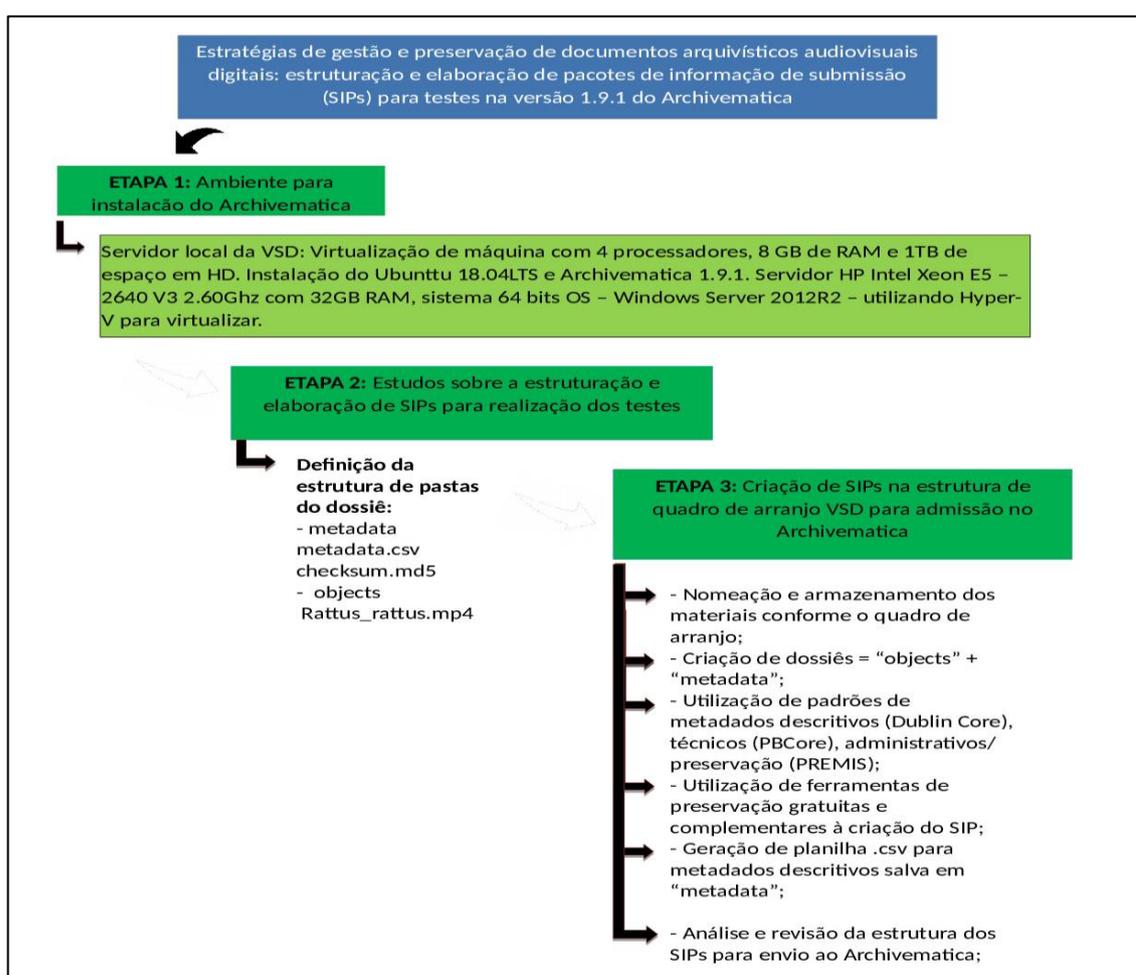
Os dados trabalhados foram submetidos ao Archivematica versão 1.9.1 com a transferência de objetos digitais produzidos pela VSD na forma de pacotes de informação de submissão (SIP) e sua posterior análise de conformação com a geração de pacotes de informação de arquivamento (AIP) pelo próprio repositório. Foram planejados e estruturados pacotes de informação elaborados dentro da

⁶ Portaria da Presidência da Fiocruz 6069-2019-PR.

estrutura de diretórios estabelecida de acordo com o quadro de arranjo da distribuidora com o objetivo de preparar previamente SIP com as características pertinentes ao processo de admissão e processamento pelo repositório.

O fluxo representado na figura 1 expõe os estudos e passos preliminares para a estruturação, elaboração e criação dos SIP testados. Posteriormente, as ações são descritas nas sessões de três etapas: Ambiente para instalação do Archivematica, Estudos sobre a estruturação e elaboração de SIP para realização dos testes e Criação de SIP na estrutura de quadro de arranjo VSD para admissão no Archivematica.

Figura 1. Estudos e passos preliminares para estruturação, elaboração e criação dos SIP (Etapas 1, 2 e 3) testados



Fonte: VídeoSaúde Distribuidora da Fiocruz, 2020.

4.1 ETAPA 1: Ambiente para instalação do Archivematica

As ações de preparação do ambiente de *hardware* para instalação do *software* no servidor local da VSD foram tratadas nesta etapa. A versão do *software* Archivematica instalada para os testes nas dependências da VSD foi à versão 1.9.1. A VSD contou com um servidor local com estrutura de diretórios baseada em quadro

de arranjo, em consonância com estrutura hierárquica da norma internacional ISAD(G). O servidor foi instalado nas próprias dependências da VSD, localizada no prédio de expansão da Fiocruz.

4.1.1 Configurações da instalação do Archivematica/VSD

Archivematica foi instalado com a seguinte configuração no servidor local da VSD: Virtualização de máquina com 4 processadores, 8 GB de RAM e 1TB de espaço em HD. Instalação do Ubuntu 18.04LTS e Archivematica 1.9.1. Servidor HP Intel Xeon E5 – 2640 V3 2.60Ghz com 32GB RAM, sistema 64 bits OS – Windows Server 2012R2 – utilizando Hyper-V para virtualizar.

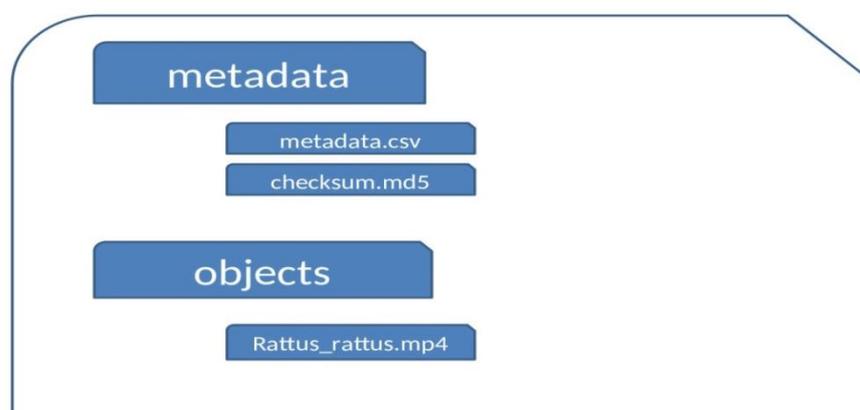
4.1.2 Fluxo para inicialização do servidor local da VSD para os testes

Sequência: Primeiro ligava-se o servidor windows com Windows 2012. Depois, rodava o virtualizador Hyper-V e posteriormente inicializava-se o OS Ubuntu 18.04LTS. Durante a inicialização era realizado um levantamento dos serviços do Archivematica para um novo acesso e retorno aos testes. Os serviços levantados no Archivematica cada vez que o servidor foi ligado para os testes estão representados na figura A-1 (anexo1).

4.2 ETAPA 2: Estudos sobre a estruturação e elaboração de SIP para realização dos testes

A etapa 2, está relacionada aos estudos sobre o conceito de pacotes de informação de submissão e o que se faz necessário para sua estruturação e elaboração visando a admissão desses SIP no *software*. Com base na estrutura de quadro de arranjo da VSD, na documentação do *software* Archivematica e no conceito de pacote informação OAIS, os pacotes de informação de submissão foram elaborados de acordo com a figura 2.

Figura 2. Exemplo de dossiê (BR_RJVSD_09_01_01_0001_01)



Fonte: VídeoSaúde Distribuidora da Fiocruz, 2020.

4.3 ETAPA 3: Criação de SIP na estrutura de quadro de arranjo VSD para admissão no Archivemática

A terceira etapa está relacionada com as ações de criação de SIP em conformidade com a estrutura de quadro de arranjo da VSD. As ações dessa etapa são descritas brevemente em subtópicos a seguir:

4.3.1 Nomeação e armazenamento dos materiais conforme o quadro de arranjo

Os SIP criados foram nomeados com base na nomenclatura estabelecida e detalhada no próximo tópico, respeitando a estrutura de arranjo da VSD.

4.3.2 Criação de dossiês = "objects" + "metadata"

Exemplo de estrutura de SIP gerado:

1) Dossiê com a nomeação correspondente ao quadro de arranjo da VSD que segue a ordem: CODEARQ+ cód. fundo+ cód. seção+ cód. série+ número do dossiê dentro da série+ cód. da forma do material (matriz, não editado, legendado, audiodescrição etc).

Ex: BR_RJVSD_09_01_01_0001_01

BR_RJVSD = CODEARQ da VideoSaúde

09 = código fundo Icict

01 = código seção VideoSaúde

01 = código da série "Programas"

0001 = número do dossiê dentro da série "Programas"

_01 = código da forma do material "Matriz"

2) Criação de duas subpastas dentro da pasta do dossiê correspondente com os seguintes nomes: "metadata" e "objects".

3) Na pasta "metadata", foram inseridas uma planilha gerada em Calc (libreoffice) e codificada para UTF-8 com metadados Dublin Core, com os rótulos das colunas de acordo com as instruções disponíveis no próprio site do Archivemática⁷ e um checksum criado (pode-se utilizar o *software* gratuito MD5Checker) para o arquivo ou conjunto de arquivos de vídeo que integrem o dossiê. A nomeação do arquivo do checksum seguiu o padrão checksum.md5.

4.3.3 Utilização de padrões de metadados descritivos (Dublin Core), técnicos (PBCore), administrativos/ preservação (PREMIS)

⁷ <https://www.archivemática.org/pt-br/docs/archivemática-1.10/>

Em seu fluxo de trabalho, a VSD faz uso dos seguintes tipos de metadados: descritivos, técnicos e de preservação. Para os metadados descritivos, é utilizado o Dublin Core, padrão amplamente aceito. Já os metadados técnicos foram extraídos por meio do *software* de código aberto MedialInfo e exportados no padrão PBCore2. Os metadados de preservação estão definidos como o PREMIS podendo haver uma complementação com outros padrões no que tange o registro de histórico de processos de digitalização de vídeos.

Os fluxos de trabalho se relacionam desde o momento de produção dos documentos até o seu arquivamento permanente e disponibilização para acesso público e aberto aos vídeos da VSD. Esses fluxos se integram a partir da captura e registro de imagens em movimento que irão compor os documentos, assim como na coleta e geração de metadados descritivos sobre o conteúdo dos mesmos, até chegarem ao estágio de ingestão via descarregamento ou transferência no servidor local para o posterior processamento técnico para arquivamento. Trata-se de um arquivamento local de SIP para posterior admissão no Archivematica. Os SIP criados somente se tornarão AIP quando forem admitidos no *software*.

Após a definição da estrutura de conteúdo dos SIP, iniciou-se a fase de admissão de conteúdo no Archivematica versão 1.9.1. Cabe uma observação aqui sobre o que tange a opção de extração de metadados técnicos. Documentos audiovisuais sejam eles analógicos ou digitais, são compostos por detalhes e informações tecnológicas embutidas na construção de seus conteúdos. Isso se reflete desde informações sobre os equipamentos que são utilizados para a sua produção até informações técnicas sobre a própria composição dos arquivos gerados como, por exemplo, formato de arquivo, taxa de bitrate, especificação de *codec*, sistema de cor, detalhes sobre a trilha de vídeo e a trilha de áudio etc.

Nesse contexto, metadados técnicos são indispensáveis aos documentos audiovisuais, pois o registro dessas informações possibilitará ao preservador desses documentos conhecer e tentar reproduzir e/ou simular as condições próximas das ideais da configuração original desse conteúdo audiovisual. Levando em consideração o exposto, decidiu-se em um primeiro momento armazenar esses metadados técnicos junto com os metadados descritivos dos SIP. Entretanto, observou-se que o metadados técnicos no ato da admissão de conteúdo no Archivematica são extraídos por default pelo mesmo *software* utilizado pela VSD na extração desses metadados, ou seja, o MedialInfo.

Constatado o fato, optou-se pela exclusão do arquivo com metadados técnicos contidos nos SIP que foram admitidos no *software*. Foi realizada uma configuração do MedialInfo do Archivematica para que fossem exibidos os metadados técnicos por ele extraídos, no formato PBCore2.

4.3.4 Utilização de ferramentas de preservação gratuitas e complementares à criação do SIP

As ferramentas (*softwares*) utilizadas para a criação dos SIP para os testes realizados são todas gratuitas e abertas, a saber: MediaInfo, Robocopy, Md5Checker, Droid e Libreoffice CALC. Para esta estratégia, tais ferramentas servem aos seguintes propósitos conforme a figura abaixo:

Figura 3. Ferramentas testadas para criação dos SIP.

FERRAMENTA	AÇÃO DE UTILIZAÇÃO	INFORMAÇÕES SOBRE
MediaInfo	Extração de metadados técnicos embutidos.	https://mediainfo.net/pt/MediaInfo
Robocopy	Verificação de integridade e manutenção de metadados técnicos.	https://docs.microsoft.com/en-us/windows-server/administration/windows-commands/robocopy
Md5Checker	Verificação de integridade e fixidez dos objetos digitais, por meio da geração de valor de <i>checksum</i> (do tipo MD5).	http://getmd5checker.com/
DROID	Descrição pormenorizada dos formatos de arquivo digitais, a fim de auxiliar no inventário.	https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/
Libreoffice CALC	Construção das planilhas de metadados descritivos, em formato .csv.	https://pt-br.libreoffice.org/descubra/calc/

Fonte: VídeoSaúde Distribuidora da Fiocruz, 2020.

Os objetos digitais audiovisuais da VSD requerem *softwares* que possuam ou comportem os *codecs*⁸ de vídeos e de áudio necessários para a leitura dos seus formatos de arquivo específicos. Como os arquivos digitais de vídeo possuem características mais complexas em comparação a outros arquivos digitais, envolvendo a presença de *codecs* específicos, esse monitoramento de *softwares* e *codecs* é um fator de grande relevância para a manutenção de longo prazo dessas informações digitais.

Após o descarregamento ou transferência dos conteúdos advindos de cartões de câmeras digitais, pendrive, FTP etc, os vídeos passaram pelo processo de verificação de vírus e tramitação por meio da ferramenta Robocopy (para manutenção dos atributos de metadados gerados no ato de gravação do “original”) para garantia de presunção de autenticidade dos conteúdos.

⁸ Codecs: É acrônimo de codificador/decodificador, dispositivo de *hardware* ou *software* que codifica/decodifica sinais.

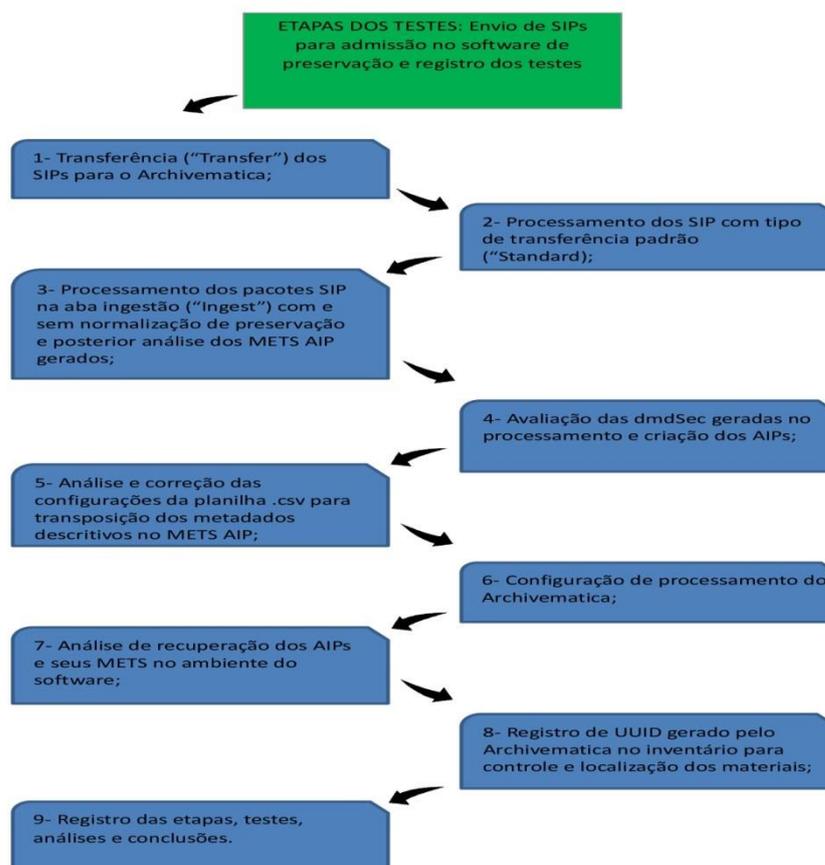
Com base na ingestão dos conteúdos, o responsável pelo tratamento técnico para criação de pacotes de informação de submissão (SIP - Submission Information Package) realizou atividades de nomeação de arquivos e armazenamento dos mesmos em dossiês criados de maneira sequencial e em conformidade com a estrutura do quadro de arranjo definido. A estrutura de dossiê foi orientada pelo modelo padrão de pasta “dossiê = título”, contendo uma subpasta “objects” – onde são depositados os arquivos a serem preservados e uma subpasta “metadata” – que abriga uma planilha (modelo anexo) com os metadados dos objetos depositados na primeira pasta. Caso haja necessidade, outras subpastas, dentro do dossiê correspondente, poderão ser criadas para atenderem a sua finalidade.

A geração e/ou extração de metadados descritivos, técnicos e de preservação, assim como o *checksum* para sua inserção nas respectivas pastas “metadata”, em seu respectivo dossiê, se deu através de ferramentas de preservação digital gratuita e de código aberto. Os metadados descritivos foram gerados em planilhas, por meio do *software* Libreoffice CALC; já os metadados técnicos embutidos foram extraídos através do MediaInfo. Finalmente os valores de *checksum* foram gerados e verificados através da ferramenta Md5Checker.

4.4 Etapas dos testes

Em meados de 2018 foram iniciados testes com o *software* Archivematica na FIOCRUZ; sistema gratuito de preservação digital de código aberto, por meio do qual determinados arquivos digitais de vídeo estão sendo admitidos a fim de se observar a eficiência da plataforma para o contexto do acervo da VSD, em particular no que se refere à sua preservação de longo prazo. Inicialmente os testes foram interrompidos por questões de infraestrutura e pelo pouco conhecimento que se tinha até então sobre o funcionamento do *software*. Apenas no segundo semestre de 2019, especificamente o período compreendido entre 09/10/2019 a 09/12/2019, é que os testes se deram de maneira robusta e continuada. No período mencionado, foram realizados 68 testes de admissão de SIP para o Archivematica da VSD. A figura 4 demonstra o fluxo das ações executadas no ato da transferência e ingestão dos SIP pelo *software*.

Figura 4. Fluxo das ações executadas no ato de transferência e ingestão de SIP



Fonte: VídeoSaúde Distribuidora da Fiocruz, 2020.

Os testes aqui descritos objetivaram planejar e elaborar a estrutura de SIP construídos fora do ambiente do *software* com vistas à sua admissão no repositório, a verificação da importação de metadados descritivos e sua transposição para as *dmdSec*⁹ correspondentes no METS¹⁰ do AIP assim como os PREMIS gerados automaticamente pelo Archivematica. Também cabe destacar que não foram trabalhados os metadados de direitos (*rights*) PREMIS nessa primeira rodada de testes.

A descrição dos testes foi realizada de maneira a agrupar em blocos os SIP que foram elaborados com características semelhantes de estrutura e composição do seu conteúdo, assim como os resultados obtidos na geração dos AIP pelo *software*. O tamanho dos arquivos testados na admissão no Archivematica estava entre 100 a 500MB, ou seja, arquivos de grande volume, entretanto, longe do volume real dos arquivos matrizes que a VSD produz (geralmente em formatos *.mov*, *.mts*, *.m2ts*, *.avi* etc).

⁹ Sigla para Descriptive Metadata Section ou Seção de Metadados Descritivos do METS.

¹⁰ Sigla para Metadata Encoding Transmission Standard ou Padrão de Codificação e Transmissão de Metadados. O esquema METS é um padrão para codificar metadados descritivos, administrativos e estruturais referentes a objetos em uma biblioteca digital, expressos usando a linguagem de esquema XML do World Wide Web Consortium (W3C). O padrão é mantido no Gabinete de Desenvolvimento de Redes e Padrões MARC da Biblioteca do Congresso Americano (LC).

A estrutura dos SIP foi testada com algumas alterações pontuais no que diz respeito ao envio de objetos e metadados contidos primeiramente nas próprias subpastas “objects” e “metadata” dos dossiês respectivamente, em um segundo momento contidos no nível da subpasta “metadata”, ou seja, sem a subpasta “objects” e, posteriormente, inseridos dentro da própria subpasta “metadata”. Os arquivos de vídeo testados estavam em formato .mov com *codec* H.264, ou seja, por default o Archivematica está preparado para trabalhar com a automatização da normalização para preservação desse formato de arquivo transcodificando através da ferramenta FFmpeg e normalizando para .mkv em FFV1.

4.4.1 Transferência dos SIP

Primeira questão a ser mencionada é sobre a transferência de conteúdo para o Archivematica. A transferência de conteúdo para o *software* só é concretizada se o terminal/computador de origem dos recursos para admissão estiver autorizado a enviar conteúdo para a área de transferência do Archivematica. Não é possível enviar arquivos isolados de uma estrutura de pasta, ou seja, arquivo (s) contido (s) numa pasta de origem para a área de transferência do *software*.

Outro ponto importante para facilitar a intensidade do ritmo de testes é a questão do ponto de acesso para envio e transferência de conteúdo direto para o *software* sem a necessidade de se trabalhar com FTP¹¹. Com isso, foi possibilitada uma celeridade na criação, atualização e envio de conteúdo para o *software*. Nesse sentido, a participação da T.I. no suporte à configuração desse arranjo é essencial.

Estabelecido esse artifício de integração direta entre o computador de realização dos testes com a área de transferência do Archivematica, os testes puderam progredir com uma dinâmica mais fluída. Por padrão, nos 68 testes realizados, os SIP foram enviados via máquina autorizada a apontar conteúdo para área de transferência (aba “Transfer”) do Archivematica e com o tipo de transferência (“Transfer type”) selecionado como a opção padrão (“Standard”).

4.4.2 Opções de normalização

Após a transferência e o processamento dos pacotes ingeridos no *software*, os mesmos são enviados para a aba “Ingest” ou admissão, onde são gerados os AIP. Nessa aba são apresentadas as opções de normalização que o *software* tem por default e que o usuário poderá optar e aprovar manualmente o tipo escolhido ou automatizar a configuração de normalização que atenda as suas necessidades específicas.

¹¹ Protocolo de Transferência de Arquivos é um protocolo padrão/genérico independente de *hardware* sobre um modo de transferir arquivos/ficheiros e também é um programa de transferência. A transferência de dados em redes de computadores envolve normalmente transferência de arquivos e acesso a sistemas de arquivos remotos. Fonte: [Wikipédia](#)

Durante os testes, a ação (manual) de optar ou não, em um primeiro momento, pela normalização para preservação desse conteúdo audiovisual, necessita de uma discussão mais aprofundada sobre a questão. Tendo em vista que, na primeira quinzena de testes com arquivos de 490MB normalizados para preservação, os mesmos chegaram aos 5.7GB com essa opção de normalização, foi observada essa demanda por uma reflexão mais técnica e detalhada. Todo esse volume gerado com a normalização de preservação terá um custo exponencial principalmente sobre a questão da capacidade de armazenamento e da velocidade de rede, que a instituição deverá arcar para garantir a manutenção da preservação desses materiais.

Observando que alguns formatos, principalmente de vídeo nato digitais, são gerados em formatos com certa consolidação e reconhecimento de seus usos (por exemplo, o formato .mov), inclusive de preservação, a sua normalização imediata pode não acarretar em ganhos de qualidade real de áudio e vídeo para o conteúdo original por ora normalizado. Isto sem falar no tamanho dos arquivos gerados no momento da normalização para preservação que podem chegar até 10x o tamanho do arquivo original. Apesar dos contêineres e *codecs* de preservação estarem em constante evolução e melhorias de implementação das recomendações de preservação digital (por exemplo o contêiner. mkv com o *codec* FFV1), isso tem um custo significativamente pesado para as instituições que lidam com acervos audiovisuais.

4.4.3 Falha na estrutura dos SIP

Esta ação tem como objetivo, avaliar se a estrutura de SIP enviado para o Archivematica foi capaz proporcionar a execução do micro serviços e seus respectivos “jobs” (ações internas de cada micro serviço) e se as informações do pacote foram transferidas e agregadas ao AIP gerado.

Nos 10 primeiros testes, foi trabalhada a estrutura de dossiê + “metadata” + “objects”. Os metadados descritivos foram gerados ainda em uma planilha Excel e salvos como .csv. As ações de processamento nesses testes seguiram o estabelecido por default do *software*. Na opção de normalização já na etapa do “Ingest”, foi acionada a opção de normalizar para preservação apenas, sem optar pela normalização de preservação e acesso. Na observação da estrutura de AIP gerados nos 10 primeiros testes, foi constatada a não transposição dos metadados enviados em Excel para o corpo de uma *dmdSec* no METS gerado. Além da codificação da planilha .csv não estar no padrão correto para interpretação do sistema, a falta da nomeação correta do arquivo de vídeo (subpasta “objects”) na coluna “filename” (obejcts/nomearquivo.extensão) também foi um fator decisivo na não criação de uma *dmdSec* no METS do AIP gerado na admissão.

Mesmo sem a transposição dos metadados descritivos, os AIP são gerados e podem ser armazenados ou não conforme decisão do usuário. Essa decisão pode ser tomada baseada no micro serviço de revisão do AIP na aba “Ingest” do Archivematica.

Ao clicar em “Review”, o usuário poderá ver a estrutura do METS que está sendo gerado e decidir se cancela o armazenamento do AIP ou se prossegue com os processos. Nessa funcionalidade, o usuário poderá ver quantas dmdSec se formaram e com isso, tomar a decisão mais indicada ao caso.

4.4.4 Codificação da planilha para exportação de metadados descritivos

Para conhecer a potencialidade da codificação da planilha .csv para exportação de metadados descritivos para transposição no METS AIP, foram realizadas consultas com instituições e usuários do Archivematica que também lidam com documentos audiovisuais (Arquivo Nacional do Brasil, Museu de Arte Moderna de Nova Iorque – MoMA, Museu da História do Computador - Computer History Museum (CHM) entre outros), além de contribuições advindas das comunidades de discussão do *software* na internet. Após ditas consultas chegou-se a uma estrutura de codificação da planilha que gera os arquivos .csv possibilitando a incorporação de metadados descritivos (sejam eles Dublin Core ou não) e geração de dmdSecs no corpo do METS AIP. Essa codificação se faz de suma importância, pois torna possível a interpretação de qualquer caractere universal padrão do Unicode, sendo também compatível com o ASCII. Foi sugerido o trabalho com o Libreoffice CALC com a seguinte codificação como demonstrada na Figura A-2 (anexo 2).

Destaca-se que, quando testadas com a utilização de vírgula na inserção da descrição dos campos DC na planilha .csv, as informações foram dispostas em campos diferentes na dmdSec que transpõe esses metadados para o corpo do METS AIP.

Até o momento dos testes não foi resolvida a questão da não interpretação de vírgulas criadas nos campos de descrição do conteúdo na planilha pela versão 1.9.1 do Archivematica. Dessa maneira, foi utilizado o ponto e vírgula (;) como substituto das vírgulas cabíveis nas sentenças de descrição.

4.4.5 Transposição dos metadados descritivos no METS AIP

Com base no que foi descrito na subseção anterior, a transposição de metadados descritivos passa a ser avaliada nos METS AIP gerados. Seguindo essas sugestões e recomendações, os testes voltaram para sua estrutura de SIP descrita na Figura 2. Os SIP foram trabalhados e analisados do teste nº 20 ao 35 de maneira a observar como se daria ou não a incorporação dos metadados descritivos dos materiais admitidos no *software*. A partir desse bloco de testes também foi utilizado um padrão de execução dos micros serviços do Archivematica visando otimizar a capacidade de processamento do *software* tendo em vista que, por trabalharmos com documentos audiovisuais nato digitais, a questão da decisão por uma normalização de preservação dos SIP enviados fica a margem do escopo desse primeiro momento

de testes. Essa questão da normalização dos formatos será objeto de uma discussão posterior.

A configuração das ações a serem tomadas manualmente no processamento dos SIP nas abas “Transfer” (transferência) e “Ingest” (admissão) foram alteradas diretamente na aba de administração do sistema (“Administration”) nas configurações de processamento (“Processing configuration”) e registradas conforme registrado nas figuras A-3, A-4 e A-5 (anexos).

Com as alterações realizadas na codificação e configuração da planilha .csv em CALC contendo os campos de metadados trabalhados no padrão Dublin Core junto ao processamento de SIP reconfigurados, foi possível chegar a uma estrutura de METS AIP contendo as dmdSecs com os metadados contidos na planilha .csv dos SIP testados, como mostram as figuras A-6 e A-7 (anexos). Abaixo, segue apresentação do trecho em XML (Figura 5) do METS AIP contendo os metadados exportados via planilha .csv e a dmdSec em Dublin Core.

Figura 5. Trecho em XML do METS AIP contendo os metadados exportados via planilha .csv e a dmdSec em Dublin Core.

```
<mets:dmdSec ID="dmdSec_2">
<mets:mdWrap MDTYPE="DC">
<mets:xmlData>
<dcterms:dublincore
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/"
xsi:schemaLocation="http://purl.org/dc/terms/
http://dublincore.org/schemas/xmls/qdc/2008/02/11/dcterms.xsd"
>
<dc:identifiser>BR_RJVSD_09_01_01_0003_01_0001de0001</dc:identifiser>
<dc:title>A história da saúde pública no Brasil - 500 anos na busca por soluções</dc:title>
<dc:creator>Fundação Oswaldo Cruz</dc:creator>
<dc:creator>Selo Fiocruz</dc:creator>
<dc:publisher>VideoSaúde - Distribuidora</dc:publisher>
<dc:date>2015</dc:date>
<dc:description>A história da saúde pública no Brasil - 500 anos na busca de soluções. Direção: Jardim Silva. Editor: Selo Fiocruz Video. Dvd (17min.): Palm; Son.; color; 2015. (citação).</dc:description>
<dc:description>Essa história começa com a chegada dos colonizadores portugueses; quando os problemas sanitários ficaram mais graves e começamos a busca de soluções para as questões de saúde dos brasileiros.</dc:description>
<dc:format>mp4</dc:format>
<dc:language>português</dc:language>
<dc:subject>saúde pública</dc:subject>
<dc:subject>história</dc:subject>
<dc:subject>SUS</dc:subject>
<dc:type>video</dc:type>
<dc:coverage>00:17mim:00</dc:coverage>
<dc:rights>open access</dc:rights>
</dcterms:dublincore>
</mets:xmlData>
```

Fonte: VídeoSaúde Distribuidora da Fiocruz (2020).

4.4.6 Inserção manual de metadados descritivos

A partir do momento em que as *dmdSec* foram geradas no corpo do arquivo METS dos AIP gerados, inicia-se a fase de teste com a inserção manual de metadados descritivos no próprio Archivematica. A inserção manual de metadados descritivos pode ser realizada na aba “Transfer”, assim que o material trabalhado esteja disponível na área de transferência do *software*, ou na aba “Ingest” quando o material está sendo processado para gerar o AIP (o *software* pode ser configurado para enviar um lembrete sobre a criação de metadados manualmente).

Essa funcionalidade se mostrou efetiva para a descrição do SIP como um todo e não é recomendada para descrever os objetos contidos no pacote de submissão. Quando o pacote SIP foi testado com mais de um objeto, a utilização da planilha .csv com a configuração já mencionada na Figura 2, se mostrou mais eficiente na medida em que tornou capaz a descrição de cada objeto representado linha por linha na planilha. Houve uma intensa preocupação com a nomenclatura correta do objeto na coluna “filename” da planilha .csv pois, como mencionado anteriormente, se nessa coluna o nome dos arquivos estiver errado, não será possível a geração de *dmdSec* que corresponda ao mesmo objeto. Com isso, ficam comprometidos e incompletos os metadados que devem compor o arquivo METS do AIP gerado.

4.4.7 Recuperação dos AIP: Acesso aos METS

Do 36º teste até o último, de número 68, realizado em 09/12/2019, os testes voltaram-se para a avaliação de recuperação dos AIP no ambiente do *software*. Também foi dada ênfase na análise e interpretação dos metadados importados via .csv (também são aceitos metadados descritivos em formato Json) e transpostos na admissão dos SIP para os METS gerados.

Os metadados PREMIS gerados pelo processamento do Archivematica foram apresentados em todas as fases de teste com as normalizações de preservação, preservação e acesso e nos casos em que optamos por não normalizar para nenhuma das finalidades citadas anteriormente. Os eventos de normalização são registrados e a eles são atribuídos UUIDs¹² no corpo do METS dos AIP gerados.

Até a versão 1.9.1 a recuperação dos arquivos METS só foi capaz de ser realizada, baixando o AIP como um todo. Essa limitação expôs uma vulnerabilidade do sistema no tocante ao gerenciamento das metainformações contidas na “alma” dos AIP, ou seja, dos arquivos METS gerados além de acarretar problemas de download caso os pacotes sejam muito grandes. Esse ponto de discussão é tema recorrente nos fóruns de discussão dos usuários do Archivematica. Outra questão a ser

¹² Um identificador único universal (do inglês universally unique identifier - UUID) é um número de 128 bits usado para identificar informações em sistemas de computação. O termo identificador único global (globally unique identifier - GUID) também é utilizado. O significado de cada bit é definido por qualquer uma das diversas variantes. (www.uuid.online)

observada é a não aplicabilidade de navegação multinível, ou seja, não é possível visualizar a estrutura hierárquica da classificação dada aos documentos.

4.4.8 Registro de UUID no inventário para controle e localização dos materiais

Com base no aumento da quantidade de AIP gerados com os testes foi observada a necessidade de gerar um inventário para controle e identificação dos AIP. Finalizando a etapa de avaliação dos AIP, foi criada uma planilha de controle externo dos materiais admitidos no *software* afim de registrar os UUIDs gerados para localização dos AIP no Archivematica como uma espécie de inventário do conteúdo admitido para o *software*.

5 Discussão e resultados dos testes

O trabalho realizado nas pesquisas que antecederam a prática nos testes com o repositório de preservação Archivematica na sua versão 1.9.1, foram fruto de um complexo conjunto de discussões internas e externas ao ambiente institucional da Fiocruz. O registro e compartilhamento dos testes visa dar suporte à prática das ações iniciais na construção de SIP com materiais audiovisuais para envio e processamento no repositório de preservação e também na verificação do conteúdo gerado pelos AIP no Archivematica.

Seguindo as orientações e recomendações já mencionadas no que tange a presunção de autenticidade e integridade de documentos arquivísticos digitais, os testes buscaram analisar como os pacotes de informação criados externamente ao Archivematica seriam processados pelo *software* e como os AIP iriam apresentar os metadados relativos à admissão desses pacotes no corpo dos METS AIP gerados. A análise se orientou na avaliação da composição dos dossiês (pacotes de informação) para admissão no *software* e a respectiva estrutura de seus componentes (pastas “objects” + “metadata”), no processamento dos SIP transferidos e na verificação dos AIP e os seus metadados METS gerados pelo Archivematica. A verificação dos METS se focou na transposição e criação de dmdSec com os metadados contidos na planilha .csv, assim como nos metadados PREMIS (PREMIS: OBJECT, PREMIS: EVENT e PREMIS: AGENT) gerados pelo micro serviços de processamento do *software*.

Muitos pontos observados durante a realização dos testes foram sendo registrados enquanto possíveis situações de tomada de decisão e objetos de futuros estudos e testes práticos na busca por soluções viáveis que garantam a preservação desse patrimônio arquivístico audiovisual em consonância com as melhores práticas estabelecidas. Questões como a decisão por normalizar para preservação os AIP gerados no momento da admissão no Archivematica, como formatar as ações específicas de processamento exclusivamente para conteúdo audiovisual que correspondam as especificidades do acervo da VSD e optar por criar pacotes de

informação de disseminação (DIP) para as plataformas de acesso¹³ da VSD entraram no hall de questões para uma discussão mais aprofundada sobre quais decisões e caminhos devem ser seguidos.

Na prática, pôde-se observar a necessidade de algumas ações complementares para o gerenciamento ativo dos AIP criados pelo *software*. A criação de uma planilha com o registro dos UUID gerados foi fundamental quando os testes ultrapassaram um limite razoável para se ter controle, ou seja, quando se pensa em trabalhar com uma grande escala de conteúdo admitido no Archivematica a utilização de um inventário com os UUIDs dos AIP pode ser necessária. Até a versão 1.9.1, o acesso aos METS AIP só era viabilizado através do download integral do AIP. Isso acarretou em mais demanda por espaço de armazenamento na máquina local destinada aos testes, na paralisação das atividades no próprio Archivematica já que este não permite outras ações acontecerem concomitantemente com o descarregamento de um AIP, além de aumentar a complexidade de gerenciamento dos AIP via acesso aos seus metadados contidos no METS.

Sobre a impossibilidade de uma navegação multinível na criação de AIP, o Archivematica até a presente versão testada 1.9.1, não apresentou nenhuma implementação de tal funcionalidade entre outras que, apesar de não estarem desenvolvidas nessa versão, são discutidas nos fóruns e eventualmente encontram-se em desenvolvimento (patrocinado ou não) já que o *software* é livre e de código aberto. Em outras palavras se você encontrar uma demanda específica para o seu trabalho de preservação com o Archivematica, você poderá optar por desenvolver a funcionalidade que supra a sua demanda, se contar com um bom suporte institucional de T.I. para isso, ou poderá contatar a empresa desenvolvedora do *software* (Artefactual Systems) para negociar pela implementação de tal funcionalidade.

O uso do Archivematica, baseado no modelo OAIS, enquanto *software* de preservação digital de documentos arquivísticos audiovisuais junto com procedimentos prévios e complementares de gestão, se mostrou uma ferramenta com um potencial cada vez mais abrangente no que se refere às ações de preservação digital capazes de dar suporte a presunção de autenticidade e integridade dos documentos audiovisuais com características semelhantes aos produzidos e testados pela VideoSaúde- Distribuidora.

6 Considerações

O uso de repositórios arquivísticos digitais confiáveis no Brasil, em consonância com a resolução nº 43 do CONARQ, para propósitos de preservação está cada vez mais difundido entre as comunidades que lidam com acervos arquivísticos sejam eles digitais ou em meio analógico aguardando processo de digitalização.

¹³ A VideoSaúde trabalha com o repositório institucional da Fiocruz baseado em DSpace (ARCA) e com o seu Banco de Recursos Audiovisuais em Saúde (BRAVS) desenvolvido internamente pelo setor de tecnologias da informação (Ctic) do Ictict.

O *software* Archivemática para fins de preservação de conteúdo audiovisual tem sido bastante difundido por instituições de referência na área de preservação no Brasil. Entretanto, pouca experiência prática de uso do repositório tem sido documentada e divulgada por parte das instituições que o adotaram enquanto ferramenta de preservação digital. A proposta do estudo objetivou contribuir de alguma maneira com o registro de ações práticas que antecederam o processo de estruturação e criação de pacotes de informação para submissão e admissão de material audiovisual no Archivemática.

Muitos avanços estão sendo alcançados e possíveis, graças ao entendimento teórico e, sobretudo prático, dos modelos de referência, das estratégias e ações de gestão e preservação de documentos arquivísticos digitais. Nesse sentido, documentar o exercício prático das ações e atividades de preservação torna-se de fundamental relevância para consolidação, atualização e inovação do conhecimento produzido em preservação digital.

Durante a realização dos testes muitas questões foram sanadas integralmente ou parcialmente tendo como base, consultas aos fóruns de discussões sobre o uso do Archivemática, incluindo suporte da equipe técnica que trabalha no desenvolvimento do *software*, bem como com especialistas nacionais e internacionais em preservação digital e audiovisual.

Concluindo, os testes registrados subsidiam e fomentam discussões institucionais e interinstitucionais acerca do uso do Archivemática na preservação digital de documentos arquivísticos audiovisuais, assim como na construção e/ou revisão de políticas e programas de preservação no âmbito da Fiocruz.

O registro dos exercícios realizados mostra a importância do compartilhamento dos conhecimentos adquiridos com as práticas de ações em preservação audiovisual digital, melhorando as possibilidades de minimizarmos erros recorrentes, de maximizar acertos e correções além de permitir uma maior abrangência e velocidade no desenvolvimento de possíveis soluções tecnológicas, que garantam que o patrimônio documental arquivístico digital seja preservado e acessado pelo tempo que for necessário.

Referências

ARQUIVO NACIONAL. CÂMARA DOS DEPUTADOS. **A preservação de documentos arquivísticos digitais**: diretrizes para organizações. INTERPARES 2 PROJECT: Diretrizes do preservador. Disponível em: http://www.interpares.org/display_file.cfm?doc=ip2_preserver_guidelines_booklet--portuguese.pdf. Acesso em: 09 ago. 2019.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT) **ABNT NBR 6023**: Informação e documentação - Referências - Elaboração: 2.ed. Rio de Janeiro: ABNT, 2018. 74 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **ABNT NBR 15472: 2007**. Sistemas espaciais de dados e informações – Modelo de referência para um sistema aberto de arquivamento de informação (SAAI), 2007.

CAVALCANTE, Lídia Eugenia. Patrimônio digital e informação: política, cultura e diversidade. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, n.23, 1º sem. 2007. Disponível em:

http://www.repositorio.ufc.br/bitstream/riufc/22732/1/2007_art_lecavalcante.pdf.

Acesso em: 22. jul 2019.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ) **Política de preservação digital**, 2019. Disponível em:

http://arquivonacional.gov.br/images/COGED/Politica_presercacao_digital.pdf.

Acesso em: 28 ago. 2019.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). Resolução nº 43, de 04 de setembro de 2015. **Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis–RDC-Arq/**, 2015. Disponível em :

https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq_diretrizes_rdc_arq_resolucao_43.pdf/view. Acesso em: 30 ago. 2019.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). Resolução nº 37, de 19 de dezembro de 2009. **Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais**. Disponível em:

http://conarq.gov.br/images/publicacoes_textos/conarq_presuncao_autenticidade_completa.pdf. Acesso em: 30 ago. 2019.

CONSULTIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS). **Reference Model for an Open Archival Information System (OAIS)**: recommended practice, issue 2, 2012. Version CCSDS 650.0-M-2. Disponível em:

<https://public.ccsds.org/pubs/650x0m2.pdf>. Acesso em: 27 ago. 2019.

EDMONDSON, Ray. **Arquivística audiovisual**: filosofia e princípios. Brasília: UNESCO, 2017.100 p.

FUNDAÇÃO OSWALDO CRUZ (FIOCRUZ). **Política de Acesso Aberto ao Conhecimento da Fiocruz**. Rio de Janeiro: Fiocruz, 2014. Disponível em:

https://portal.fiocruz.br/sites/portal.fiocruz.br/files/documentos/portaria_-_politica_de_acesso_aberto_ao_conhecimento_na_fiocruz.pdf. Acesso em: 27 ago. 2019.

FUNDAÇÃO OSWALDO CRUZ (FIOCRUZ). **Política de preservação dos acervos científicos e culturais da Fiocruz**. Rio de Janeiro: Fiocruz, 2018. Disponível em:

<https://portal.fiocruz.br/documento/politica-de-preservacao-dos-acervos-cientificos-e-culturais-da-fiocruz-0>. Acesso em: 27 ago. 2019.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 14721: 2012**: space data and information transfer systems – Open Archival Information System (OAIS) – Reference model, 2012.

LIBRARY OF CONGRESS. **METS**: an overview & tutorial. Sept. 2011. Disponível em: <http://www.loc.gov/standards/mets/METSOverview.v2.html>. Acesso em: 13 set. 2019.

LIBRARY OF CONGRESS. **PREMIS**: Preservation Metadata Maintenance Activity. 2015. Disponível em: <http://www.loc.gov/standards/premis/>. Acesso em: 17 set. 2019.

MÁRDERO ARELLANO, Miguel Ángel. **Critérios para a preservação digital da informação científica**. 354 p. Tese (Doutorado em Ciência da Informação). Universidade Federal de Brasília, Departamento de Ciência da Informação, 2008. Disponível em: http://bdtb.bce.unb.br/tesesimplificado/tde_busca/arquivo.php?codArquivo=4547. Acesso em: 18 set. 2019.

SANTOS, Henrique Machado; FLORES, Daniel. Preservação de documentos arquivísticos digitais: reflexões sobre as estratégias de migração. **Prisma.com**. n. 37, p. 42-54, 2018. Disponível em: <https://pentaho.letras.up.pt/index.php/prisma.com/article/viewFile/4707/4395>. Acesso em: 15 jul. 2019.

SARAMAGO, Maria de Lurdes. Metadados para a preservação digital e aplicação do Modelo OAIS. **Anais...** In: CONGRESSO NACIONAL DE BIBLIOTECARIOS, ARQUIVISTAS E DOCUMENTALISTAS, 8. 2004, Estoril. Disponível em: <https://www.bad.pt/publicacoes/index.php/congressosbad/article/view/640>. Acesso em: 01 out. 2019.

SAYÃO, Luís Fernando. Uma outra face dos metadados: informações para a gestão da preservação digital. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 15, n. 30, p.1-31, 2010. Disponível em: <https://www.icict.fiocruz.br/sites/www.icict.fiocruz.br/files/Uma%20outra%20face%20dos%20metadados.pdf>.

SILVA, R. R. G. da; et. Al A iniciativa Legatum e a preservação digital de arquivos audiovisuais públicos. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 14, n. 3, p. 515 - 540, set. 2016. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/8646279>. Acesso em: 22 jul. 2019.

VÍDEOSAÚDE Distribuidora da Fiocruz (VSD). **Ferramentas testadas para criação dos SIP**. Rio de Janeiro: Videosaúde - Distribuidora da Fiocruz, 2020.