

# PERFIL DE PADRÕES DE METADADOS DE PRESERVAÇÃO PARA DOCUMENTOS ARQUIVÍSTICOS AUDIOVISUAIS

## PRESERVATION METADATA STANDARDS PROFILE FOR AUDIOVISUAL ARCHIVAL DOCUMENTS



### Resumo

**Introdução:** Este artigo visa apresentar uma proposição de um perfil de padrões de metadados que atenda a requisitos preservação digital e as especificidades dos documentos arquivísticos audiovisuais videográficos da VideoSaúde – Distribuidora (VSD), serviço do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT) da Fiocruz. **Objetivo:** O artigo tem como objetivo propor a formulação de um perfil de padrões de metadados que atenda a requisitos de preservação digital pautados em princípios arquivísticos e orientados pelo modelo de informação proposto na norma ISO 14721-2012 (OAIS). **Metodologia:** A metodologia aplicada na elaboração do artigo apresenta uma abordagem amparada no caráter multidisciplinar de uma pesquisa do tipo exploratória e descritiva, visto o reconhecimento de uma ampla contribuição advinda de diversas áreas do conhecimento para com a temática. **Resultados:** Os padrões de metadados apontados para a definição do perfil adotado mostraram considerável robustez na composição das Informações de Descrição de Preservação (IDP ou PDI) necessárias para compor a estrutura de Pacotes de Informações de Arquivamento (AIP) OAIS presumidamente autênticos, confiáveis e interoperáveis. **Conclusão:** O artigo concluiu que estudos sobre a aplicabilidade de metadados – e as tecnologias que envolvem os seus usos – devam ser programados e atualizados enquanto estratégia fundamental para ações de preservação e acesso a longo prazo ao acervo arquivístico audiovisual da VSD, além de propiciarem sustentabilidade efetiva e eficiente ao Programa de Preservação Digital da Fiocruz.

**Palavras-chave:** Patrimônio cultural; Audiovisual; Preservação digital; OAIS; Metadados.

### Abstract

**Introduction:** This article aims to present a proposal of a metadata standards profile that meets the digital preservation requirements and the specificities of audiovisual videographic archival documents of VideoSaúde - Distribuidora (VSD), service of the Institute of Communication and Scientific and Technological Information in Health (ICICT) of Fiocruz. **Objective:** The article aims to propose the formulation of a metadata standards profile that meets the requirements of digital preservation based on archival principles and guided by the information model proposed in ISO 14721-2012 (OAIS). **Methodology:** The methodology applied in the preparation of the article presents an approach supported by the multidisciplinary nature of an exploratory and descriptive research, given the recognition of a broad contribution coming from various areas of knowledge to the theme. **Results:** The metadata standards pointed out for the adopted profile definition showed considerable robustness in the composition of the Preservation Description Information (IDP or PDI) necessary to compose the structure of Archival Information Packages (AIP) OAIS presumed authentic, reliable and interoperable. **Conclusion:** The article concluded that studies on the applicability of metadata - and the technologies surrounding its uses - should be programmed and updated as a fundamental strategy for preservation actions and long-term access to the VSD's audiovisual archival collection, in addition to providing effective and efficient sustainability to Fiocruz's Digital Preservation Program.

**Keywords:** Cultural heritage; Audiovisual; Digital preservation; OAIS; Metadata.



João Guilherme Nogueira Machado

Fundação Oswaldo Cruz

E-mail: [guilherme.machado@icict.fiocruz.br](mailto:guilherme.machado@icict.fiocruz.br)

Rio de Janeiro – RJ – Brasil



#### LICENÇA DE USO

Os autores cedem à [Revista Brasileira de Preservação Digital](#) os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a Licença Creative Commons Attribution (CC BY) 4.0 International. Esta licença permite que terceiros remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

#### PUBLISHERS

Universidade Estadual de Campinas – Sistema de Bibliotecas / Instituto Brasileiro de Informação em Ciência e Tecnologia – Rede Brasileira de Serviços de Preservação Digital – Cariniana. As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

#### EDITORES

Gildenir Carolino Santos, Miguel Angel Márdero Arellano.

#### CREDIT

**RECONHECIMENTOS:** Não aplicável.

**FINANCIAMENTO:** Não aplicável.

**CONFLITOS DE INTERESSE:** O autor certifica que não têm interesse comercial ou associativo que represente um conflito de interesses em relação ao manuscrito.

**APROVAÇÃO ÉTICA:** Não aplicável.

**DISPONIBILIDADE DE DADOS E MATERIAL:** Não aplicável.

**CONTRIBUIÇÕES DOS AUTORES:** Conceituação, Curadoria de Dados, Análise Formal, Aquisição de Financiamento, Investigação, Metodologia, Administração de Projetos, Recursos, Software, Supervisão, Validação, Visualização, Redação – rascunho original; Redação – revisão & edição: MACHADO, J. G. N.

Submetido em: 30/04/2023 – Aceito em: 19/05/2023 – Publicado em: 10/06/2023

## 1 INTRODUÇÃO

As novas configurações de criação e disseminação da informação potencializadas pelos avanços tecnológicos das primeiras décadas do século XXI apontam para possíveis mudanças no âmbito teórico-metodológico no campo do patrimônio cultural, em especial, no universo dos arquivos.

Os “arquivos do presente” – não mais do futuro – se deparam cada vez mais com novos suportes e linguagens utilizadas para registrar informações.

Um dos produtos dessa fase, os documentos audiovisuais, caracterizam-se de fato, como documentos de arquivo, isto é, aqueles recebidos e/ou produzidos por pessoas ou entidades públicas ou privadas no exercício de suas atividades. Esse gênero documental, desde seu advento, tem encontrado dificuldades para ser incorporado às instituições arquivísticas e tratado dentro dos padrões arquivísticos de organização documental. Desse modo, os documentos audiovisuais têm sido uma problemática constante dentro dos arquivos, uma vez que os profissionais e, a própria literatura, não têm dado conta de suprir essa necessidade. (SANTANA e MADIO, 2013, p. 36)

Compreender essas novas configurações e dinâmicas dos acervos arquivísticos – parte do patrimônio cultural – frente aos avanços tecnológicos se faz cada vez mais eminente e necessário, possibilitando atualizações, adaptações e/ou correções teórico-metodológicas que, de fato, se fazem necessárias à gestão e preservação desse patrimônio e apresentam-se enquanto uma demanda na era da informação digital.

Como um dos gêneros documentais mais impactados com questões tecnológicas como a obsolescência de equipamentos, suportes, formatos e pela alta demanda por capacidade de armazenamento, o audiovisual também se destaca por ser um poderoso instrumento de comunicação e compartilhamento de conhecimento, com uso cada vez mais exponencial devido às novas tecnologias de comunicação e informação (TIC), especialmente com o advento da Internet.

A Fiocruz é uma das mais proeminentes instituições de pesquisa em saúde pública do Brasil e o maior fabricante e fornecedor de vacinas para o Sistema Único de Saúde (SUS) do país. Instituição pública estratégica para o estado brasileiro e de excelência na área de pesquisa, a Fiocruz trabalha enquanto um agente fundamental na criação, execução e inovação de políticas públicas em saúde.

A comunicação e a informação científica e tecnológica em saúde são entendidas e tratadas pela instituição como um direito humano incondicional para garantir o desenvolvimento, o progresso e reduzir as iniquidades existentes na sociedade brasileira. Nesse sentido, o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict) da Fiocruz, credita ao audiovisual um papel central na disseminação da informação, além de potencial contributo para a memória da saúde pública brasileira.

Através de políticas institucionalizadas, a Fiocruz vem desenvolvendo estudos e estratégias de preservação sob os auspícios do PRESERVO<sup>1</sup> – coordenado pela Casa de Oswaldo Cruz (COC) - com o objetivo de consolidar as diretrizes e objetivos contidos em suas políticas: difundir o conhecimento produzido pela instituição, preservando e dando acesso (aberto) por longo prazo ao seu patrimônio científico e cultural em saúde.

Os usos cada vez mais demandados ao audiovisual dentro da própria 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. É, portanto, inegável o poder informativo e de registro histórico e cultural contido intrinsecamente na linguagem audiovisual.

Dado o valor que corresponda ao audiovisual,

Os documentos audiovisuais são tão importantes – em alguns casos mesmo mais importantes – quanto outros tipos de documentos e artefatos. A relativa novidade de sua invenção, seu caráter frequentemente popular e sua vulnerabilidade às rápidas mudanças tecnológicas não diminuem sua importância. Sua conservação e o acesso a eles devem ser garantidos, assim como os recursos necessários para isso. (EDMONDSON, 2017, p. 9)

Nesse contexto, a preservação da história científica e cultural da saúde brasileira registrada em vídeo sob custódia da Fiocruz torna-se um potencial ativo de valorização da memória da instituição.

Base empírica para a pesquisa e com um histórico de 34 anos de existência, a VSD custodia um significativo acervo audiovisual arquivístico 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.

Enxergando o novo cenário como uma realidade cotidiana e focal no que tange o seu objeto principal de atividades, ou seja, a produção de documentos arquivísticos videográficos, a VideoSaúde – Distribuidora da Fiocruz (VSD) vinculada ao Ict, em parceria com outras instâncias e institutos da fundação, se aprofundou em estudos no âmbito da preservação audiovisual baseada em requisitos arquivísticos.

Com a evolução das atividades exercidas pela VSD no âmbito da Fiocruz e, conseqüentemente, no que tange à formação do seu acervo, a dimensão das atividades de produção de conteúdo audiovisual reforça ainda mais o caráter de documento de arquivo dos vídeos que compõe parte do seu acervo e que foi o foco da pesquisa.

<sup>1</sup> Complexo de Acervos da Fiocruz, com a finalidade de preservar, valorizar e divulgar o patrimônio histórico, cultural e científico institucional e contribuir para a valorização da memória das áreas de Ciências da Saúde e Biomédicas.

Entendendo o audiovisual como um gênero documental de arquivo, dotado de uma linguagem comunicacional e suporte físico de registro das informações, a VSD realiza atualizações metodológicas de tratamento técnico aplicado ao seu conjunto documental videográfico enquanto documentos arquivísticos.

Nesse sentido a VSD entende que é imprescindível que o tratamento arquivístico adequado seja aplicado aos seus documentos videográficos e que a preservação é fundamental para a existência e manutenção de acervos audiovisuais devido às especificidades e variedades de suportes, formatos e meios de armazenamento de conteúdos analógico e digital e suas possíveis instanciações<sup>2</sup>.

A complexidade das estratégias, ações e procedimentos técnicos voltados para o gerenciamento ativo de arquivos de imagens em movimento faz com que a revisão e atualização de saberes e conhecimentos teórico-metodológicos voltados para a preservação, sobretudo a preservação digital deste gênero, se construa de maneira mais dinâmica, multidisciplinar e veloz do que qualquer outro gênero documental em arquivos fazendo com que a preservação de longo prazo de conteúdo audiovisual digital seja envolta por desafios particulares e bem específicos deste gênero documental.

No caso da VideoSaúde, essa realidade é refletida através de algumas questões particulares ligadas à gestão da preservação do seu acervo. Baseadas em fatores citados anteriormente, essas particularidades se relacionam desde questões de ordem orçamentária, de recursos humanos especializados em preservação digital, de infraestrutura para armazenamento de conteúdo nascido digital, suporte de T.I., até o seu volumoso acervo em meio analógico registrado em diversos formatos de fitas de vídeo que já se encontram em total ou parcial estado de obsolescência. Assim sendo, faz-se necessária e emergencial a priorização de ações voltadas para amenizar os impactos indesejados desse contexto.

Para além das questões já mencionadas, ao contextualizar o audiovisual enquanto documento arquivístico, esse artigo busca abordar o entendimento das demandas de metainformações que atendam as qualidades de confiabilidade, autenticidade e interoperabilidade necessárias especificamente para documentos geridos e preservados por sistemas de informações digitais, sobretudo em repositórios digitais.

O artigo buscou se apoiar em recentes e consolidadas teorias e práticas dos usos de padrões de metadados que proporcionem, especialmente à VideoSaúde e a possíveis produtores de conteúdo audiovisual na Fiocruz, a elaboração estrutural de Pacotes de Informação preservados em repositórios digitais (de preservação e/ou acesso) baseados nas recomendações do modelo de referência Open Archival Information System – OAIS (ISO, 2012) e suas diretrizes de preservação por longo

<sup>2</sup> 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.

prazo de materiais digitais com ênfase à presunção de autenticidade e à interoperabilidade dos pacotes.

O OAIS é um modelo de referência conceitual e de alto nível para um Sistema Aberto de Arquivamento de Informações ou SAAI e se desdobra na descrição de um modelo de entidades funcionais e um modelo de informação. O artigo deu enfoque ao modelo de informação e seus principais conceitos e terminologias segunda a norma.

O conceito de Pacotes de Informação é peça central para o desenvolvimento e perspectivas apresentados e objetiva trabalhar com a questão das categorias de informações necessárias para comporem as Preservation Information Description (PDI) ou Informações de Descrição de Preservação (IDP) mais especificamente na construção dos Archival Information Packages (AIP) ou Pacotes de Informações de Arquivamento (AIP). Os AIP, o objeto central da preservação segundo o OAIS, devem contemplar 5 categorias de informações de PDI para que esses pacotes possam ser considerados presumidamente autênticos e confiáveis. As categorias de PDI são: referência, proveniência, contexto, fixidez e direitos.

Disto isto, o artigo se propôs a reunir, entender, organizar e apresentar um conjunto de padrões de metadados que contemplem requisitos de preservação baseados nas recomendações do modelo conceitual de referência denominado Open Archival Information System (OAIS) e que atendam às especificidades do gênero documental audiovisual produzido no âmbito da Fundação Oswaldo Cruz (Fiocruz).

O trabalho aqui proposto buscou compreender e agregar estudos mais consolidados sobre padrões de metadados<sup>3</sup> capazes de se integrarem e/ou complementarem no sentido de proporcionar uma maior completude de metainformações necessárias à gestão da preservação de documentos arquivísticos videográficos digitais. Para que as ações de geração e gerenciamento desse conjunto de metadados de maneira independente, foi necessário o uso de ferramentas específicas e, preferencialmente, gratuitas e de código aberto.

## 2 REVISÃO DA LITERATURA

A crescente dependência da sociedade em relação às informações geradas em meio digital reflete uma preocupação que, nos dias atuais, vem ganhando cada vez mais força. Inicialmente, as práticas relacionadas com a preservação digital baseavam-se na ideia de garantir a longevidade de acervos documentais. Entretanto, atualmente a preocupação é comprometida, em grande parte, pelo desconhecimento – que parece predominar em profissionais envolvidos com a custódia e gestão de acervos – sobre estratégias de preservação digital e sobre o arcabouço teórico que deve reger suas práticas.

<sup>3</sup> Dados estruturados que descrevem e permitem encontrar, gerenciar, compreender e/ou preservar documentos arquivísticos ao longo do tempo. Disponível em: <https://tinyurl.com/mrx5k7rz>.



Segundo Márdero Arellano (2004):

Com o aumento da produção de informação em formato digital, tem sido questionada cada vez mais a importância de se ter garantida a sua disponibilização e preservação por longos períodos de tempo. Essa preocupação envolve tanto os produtores dos dados quanto os órgãos detentores dessa informação. (MARDERO ARELLANO, 2004, p.16)

Ainda segundo esse autor, torna-se cada vez mais premente, para detentores de acervos digitais, a necessidade de contar com mecanismos que garantam a preservação de seus documentos em formato digital. Essa preocupação ganha destaque entre comunidades responsáveis em gerir arquivos, bibliotecas, museus e outros centro de pesquisa e informação, para os quais o desenvolvimento de padrões e o amparo legal para lidar com documentos digitais precisam de estratégias metodológicas (instituídas através de políticas, programas e planos institucionais) bem definidas.

Enquanto a norma internacional de maior destaque no campo da preservação digital, o modelo de referência OAIS para preservação de informações digitais recomenda atribuir a estas informações, características que as tornem suficientemente completas, com o objetivo de permitir que sejam interpretadas, compreendidas e utilizadas por seus usuários sem ter que recorrer a recursos especiais não amplamente disponíveis, incluindo indivíduos. A estratégia de preservação capaz de dar conta dessa atribuição é a adoção de padrões de metadados voltados para a preservação digital.

Metadados podem ser entendidos como um conjunto de dados que descreve e fornece informações sobre outros dados. Na Ciência da Informação, o seu uso, muitas vezes, é visto como uma solução para descrição e representação de objetos digitais em ambientes informacionais visando a sua recuperação de maneira mais ágil e acurada. Entretanto, no universo da preservação digital e no escopo mais amplo dos seus usos, propósitos e funções específicas, o conceito de metadados pode ser entendido de acordo com as definições a seguir.

Segundo Alves (2010, p.47), metadados são:

[...] atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades [...] Os padrões de metadados são estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados.

Para Grácio (2012, p.121), metadados podem ser definidos como “[...] um conjunto de dados, chamados ‘elementos’, cujo número varia de acordo com o padrão adotado, e que descreve o recurso, possibilitando a um usuário ou a um mecanismo de busca acessar e recuperar esse recurso”.

A partir das definições apresentadas, pode-se observar que os conceitos de metadados foram ampliados e estabelecidos para atender a determinadas finalidades e funções de acordo com as demandas das comunidades-alvo usuárias dessas

informações, como é o caso dos padrões de metadados para preservação digital. Em consonância com objetivo do artigo, serão apresentados alguns conceitos de metadados de preservação segundo alguns autores de destaque na área de preservação digital.

Sayão (2010, p.10) conceitua os metadados de preservação da seguinte maneira: “[...] de uma forma simples e direta, como a informação que apoia e documenta a preservação de longo prazo de materiais digitais”. Saramago (2004, p.1) aponta para seguinte abordagem sobre metadados de preservação: “[...] como informação de apoio aos processos associados com a preservação digital de longo prazo”. Para Arellano (2008), os mesmos estariam associados ao conteúdo dos materiais digitais, seu contexto e estrutura de criação/produção, além das alterações ocorridas em todo o seu ciclo de vida.

Os metadados podem incluir uma grande variedade de informações, e comunidades distintas têm usos diferenciados para a metaformação. Muitas vezes existem variados padrões de metadados necessários para diferentes propósitos. Devem conter informação técnica e administrativa sobre as decisões e ações preservacionistas adotadas, documentar os efeitos das estratégias de conversão de dados, assegurar a autenticidade dos recursos digitais ao longo do tempo, registrar informação acerca de gestão de coleções e de direitos e, ainda, fornecer informação acerca dos próprios metadados (SARAMAGO, 2004).

Diante dessa perspectiva, Sayão aponta que

Na medida em que a ideia de metadados se torna uma parte essencial do mundo digital, eles se mostram conceitualmente mais complexos e mais abrangentes, apoiando um espectro extremamente amplo de atividades. Essas novas dimensões de metadados são vitais para o acesso e para a interpretação dos recursos informacionais digitais; como são importantes também para a estruturação e para os processos de gestão associados a esses recursos, que podem incluir inúmeras funções, tais como: controle dos direitos, intercâmbio, comércio eletrônico, interoperabilidade técnica e semântica, reuso da informação e curadoria digital, para citar alguns (SAYÃO, 2010, p.3)

Assim, tem se tornado um consenso na área de preservação digital que os metadados de preservação devam abarcar elementos de padrões específicos para cada finalidade de uso (descrição, acesso, gestão da preservação etc) das distintas tipologias de acervo (biblioteconômico, arquivístico, museológico etc) e de gêneros documentais (audiovisuais, textuais, iconográficos, sonoros etc). Em outras palavras, metadados descritivos, técnicos, de direitos e administrativos compõem de maneira integrada e não excludente o que se entende por metadados de preservação.

Desta forma, levando em consideração a ampla gama de funções que devem ser cumpridas na execução da preservação digital, torna-se notória a complexa tarefa de estabelecer padrões de metadados que contemplem requisitos de preservação baseados nas recomendações do modelo conceitual OAIS e que atendam às especificidades dos documentos arquivísticos audiovisuais.



Especificamente no caso da VSD, a proposição de um perfil de padrões de metadados levou em conta a utilização de ferramentas de código aberto e gratuitas no suporte a construção das PDI de Pacotes SIP para sua posterior admissão no Repositório Digital Confiável (RDC) de Preservação adotado pela Fiocruz, ou seja, o Archivematica.

Com a testagem do Archivematica pela VSD para a admissão de Pacotes SIP, foi observado sucesso com a geração de Pacotes AIP robustos e completos em suas PDI. Vale ressaltar que, para documentos oriundos de processos de digitalização, a VSD vem realizando outros estudos sobre padrões de metadados específicos para audiovisuais de cunho arquivístico, sobretudo, para o desenvolvimento de uma plataforma de controle e registro do histórico de processo de digitalização enquanto evidência para a presunção de autenticidade dos representantes digitais em convergência com o exigido para o uso do RDC de Preservação Fiocruz.

### 3 MATERIAIS E MÉTODOS

O Programa de Preservação Digital (PPD) da Fiocruz estabeleceu a norma ISO/OAIS como o modelo de referência a ser adotado pela instituição no que se refere a definição da arquitetura de dados e dos aspectos operacionais dos sistemas de gestão da preservação digital dos seus acervos científicos e culturais.

Em sua seção de nº 7 “SISTEMAS DE GESTÃO”, o PPD Fiocruz descreve a estrutura de um repositório de preservação o qual deve estar baseado no modelo de entidades funcionais proposto pelo OAIS. Nesse sentido, a orientação converge com a resolução nº43 do CONARQ que estabelece diretrizes para a implementação de repositórios digitais confiáveis para a transferência e recolhimento de documentos arquivísticos digitais.

Atendendo os requisitos de preservação digital estabelecidos pelo OAIS, em especial, o uso de padrões de metadados para registrar as diferentes categorias de metainformação que devem compor pacotes AIP, demonstra-se necessária a testagem de ferramentas capazes de gerar, editar, extrair e exportar esses diferentes tipos de metadados.

Segundo Dobрева, Kim e Ross (2007), os metadados de preservação estão no centro das atividades que garantem a sustentabilidade e usabilidade dos documentos digitais em longo prazo, mas a geração automatizada dos metadados não está muito bem integrada aos fluxos de trabalho. Os autores ainda enfatizam que dentre os principais desafios para um registro de qualidade de metadados estão: a motivação, a dificuldade com a aplicação, dificuldade de compreensão do escopo dos padrões e de seu uso e a subsequente recuperação de informações.

O PPD/VSD (FIOCRUZ, 2020), objetiva apontar estratégias de preservação a longo prazo para os documentos audiovisuais digitais de seu acervo. Dentre as

estratégias apresentadas está um fluxo de trabalho que inclui o uso de ferramentas específicas para execução de operações técnicas que visam assegurar a presunção de autenticidade e integridade dos Pacotes de Informações que devem ser preservados.

As ferramentas testadas<sup>4</sup> foram analisadas sob o viés da sua capacidade de geração, extração, interoperabilidade e exportação/importação de metainformação trabalhada a partir do perfil de padrões estabelecidos na VSD e que subsidiou, sobretudo, a elaboração de Pacotes de Informação SIP OAIS para admissão e geração de AIP e as suas respectivas PDI através do Archivematica.

## 4 DISCUSSÃO E RESULTADOS

O consenso sobre os tipos de metadados e suas funções abordado para efeitos desse artigo, aponta para a definição de 3 (três) tipos mais abrangentes dessas metainformações.

Sayão (2010, p.5) e Saramago (2004), apontam para 3 (três) tipos de metadados:

- Metadados descritivos: é a face mais conhecida dos metadados, são eles que descrevem um recurso com o propósito de descoberta e identificação; podem incluir elementos tais como título, autor, resumo, palavras-chave e identificador persistente.
- Metadados estruturais: são informações que documentam como os recursos complexos, compostos por vários elementos, devem ser recompostos e ordenados. Por exemplo, como as páginas de um livro, digitalizadas separadamente, são vinculadas entre si e ordenadas para formar um capítulo.
- Metadados administrativos: fornecem informações que apoiam os processos de gestão do ciclo de vida dos recursos informacionais. Incluem, por exemplo, informações sobre como e quando o recurso foi criado e a razão da sua criação. Nessa categoria, estão metadados técnicos que explicitam as especificidades e dependências técnicas do recurso; inclui também os metadados voltados para apoio à gestão dos direitos relacionados ao recurso.

O uso de metadados para a preservação de documentos arquivísticos digitais, exige uma variada gama de padrões/ esquemas de metainformação capazes de abordar as distintas demandas dos diferentes gêneros documentais (audiovisuais, iconográficos, sonoros, textual, etc).

No campo do patrimônio cultural, os profissionais da informação ligados ao processamento técnico de acervos arquivísticos, bibliográficos, museológicos etc,

<sup>4</sup> Planilhas CSV produzidas pelo software Calc LibreOffice; MedialInfo, Bagger 2.8.1 e Archivematica.

aplicam o conceito de metadados às informações que eles utilizam na organização, descrição, identificação e acesso aos documentos contidos em suas coleções, fundos, etc. Essa metainformação deve ser regida pela utilização de padrões e práticas desenvolvidas pelas comunidades profissionais correspondentes a cada tipologia de acervo. Essa padronização visa garantir uma maior qualidade, consistência e interoperabilidade entre os sistemas de informação que gerenciam a preservação e o acesso a esse patrimônio.

O conceito de padrão de metadados pode bem visualizado na definição dada pelo IBICT que traz a seguinte abordagem:

Os metadados são agrupados em esquemas, que os organizam, normalizam e os descrevem, criando padrões. Esses esquemas permitem o melhor entendimento da finalidade de cada metadado e são, geralmente, mantidos por instituição ou organização, permitindo assim o melhor uso e, principalmente, a troca de informação entre iniciativas que utilizam o mesmo esquema de metadados.

Existem vários esquemas de metadados [...] cada qual com as suas particularidades e mais adequado a uma finalidade específica. Assim, com o objetivo de cobrir todos os aspectos importantes de uma informação em um suporte, para uma determinada finalidade, cria-se esquemas de metadados que contenham campos mais apropriados. (IBICT WIKI, online, 2013)

O surgimento e a expansão do uso da Internet foram acompanhados por uma proliferação de esquemas de metadados, cada qual com suas especificidades e funções, projetados e desenvolvidos com base em requisitos de comunidades de usuários específicos, usuários pretendidos, tipos de materiais, domínios de assuntos, necessidades de projetos etc. (CHAN; ZENG, 2006)

Ainda segunda as autoras, os termos “esquema” e “conjunto de elementos” (element set) tem sido usado para se referir a padrão de metadados. Um conjunto de elementos de metadados deve possuir dois componentes básicos:

1. Semântica – definições dos significados dos elementos e seus refinamentos.
2. Conteúdo – declarações ou instruções de quais e como os valores devem ser atribuídos aos elementos.

Nesse sentido, o uso de vocabulários controlados é essencial para busca de uma entrada de informações padronizada para cada elemento pertinente e definido no “elemento set”. Esses vocabulários são de responsabilidade das iniciativas, instituições e organizações desenvolvedoras de padrões, assim como pelo próprio gerador e administrador do conjunto de elementos adotado, desde que haja uma documentação sistematizada e atualizada sobre os termos e conceitos indicados para o seu uso.

Um esquema ou padrão de metadados deve possuir os dois componentes citados anteriormente mais a codificação do seu conjunto de elementos com uma

linguagem de marcação do tipo XML (Extensible Markup Language) a qual efetiva a capacidade de uma interoperabilidade entre sistemas de informação/ repositórios.

A seguir, alguns padrões de metadados mais utilizados por repositórios digitais no âmbito da preservação digital são apresentados de maneira mais sintética e geral, não abarcando uma análise pormenorizada dos seus conjuntos de elementos. Os padrões discutidos foram delimitados pelo escopo de uso arquivístico em repositórios digitais destinados a processar, armazenar e disseminar documentos de arquivo de guarda permanente. Foram apontados os pontos mais gerais das especificidades intrínsecas ao padrão elencado.

Relacionado à catalogação e descoberta de recursos na Internet, o padrão mais popularmente conhecido e utilizado é o Dublin Core (DC). O Dublin Core Metadata Element Set (DCMES) surgiu de uma reunião de 1995 em Dublin, Ohio, que teve como foco, metadados para informações eletrônicas em rede. Os participantes foram incumbidos de identificar um conjunto central de características comuns à maioria dos tipos de informação digital. (SAYÃO, 2010; RILEY, 2017)

É um padrão de metadados utilizado em softwares para repositórios de preservação e de acesso como o Archivematica, DSpace, AtoM, Preservica etc.

O padrão Encoded Archival Description (EAD) foi desenvolvido através de um projeto da Berkeley Library da Universidade da Califórnia e é mantido pelo Subcomitê Técnico para Padrões de Arquivamento Codificado da Sociedade de Arquivistas Americanos, em parceria com a Biblioteca do Congresso. O padrão é codificado em XML e vem sendo cada vez mais utilizado por arquivistas do mundo inteiro. (SACRAMENTO, BAIÃO, FERREIRA, 2015)

O desenvolvimento do esquema foi sendo possibilitado ao longo do tempo devido à uma gama de esforços colaborativos entre vários arquivistas e repositórios que enxergaram a utilidade do EAD para a descrição e compartilhamento de informações arquivísticas detalhadas e em conformidade com a Norma Internacional de Descrição Arquivística (ISAD-G). Por possibilitar uma estruturação multinível de metainformação baseada na ISAD-G, o EAD se torna um padrão reconhecidamente apropriado para descrição de documentos arquivísticos digitais. (LIBRARY OF CONGRESS, 2019).

Os padrões DC e EAD estão mais relacionados com a descrição, descoberta e recuperação de documentos digitais e não especificamente com a preservação digital de longo prazo (FORMENTON, CASTRO, GRACIOSO, et al., 2017). Entretanto, tais padrões são capazes de registrar uma gama de informações que têm uma contextualização de uso e suporte às categorias de PDI de Pacotes de Informações OAIS.

No campo das instituições de patrimônio, entre os padrões de metadados que conjuguem atributos técnicos para imagens digitais, um padrão em especial recebe a chancela indicativa de aplicação do Conselho Internacional de Arquivos (ICA), da

National Information Standards Organization (NISO, EUA), Library of Congress (LOC, EUA) e da Online Computer Library Center (OCLC). O ANSI/NISO Z39.87-2006 ou Metadata for Images in XML (MIX) é um padrão de que define um conjunto de elementos de metainformação voltado para imagens digitais raster para permitir que usuários desenvolvam, troquem e interpretem arquivos de imagens digitais. Foi elaborado com o objetivo de permitir a interoperabilidade entre sistemas, softwares e apoiar a gestão e o Acesso em: longo prazo as coleções de imagens digitais (NISO, 2017).

O padrão mais reconhecido de metadados de preservação é o PREMIS (Preservation Metadata: Implementation Strategies). O dicionário PREMIS e o modelo OAIS convergem para um entendimento sobre as metainformações necessárias para a gestão da preservação digital em longo prazo.

No âmbito dos repositórios digitais utilizados por instituições de patrimônio cultural, o padrão de metadados estruturais mais utilizado é o Metadata Encoding and Transmission Standard ou METS (NISO, 2008). O METS suporta uma variedade de padrões distintos de metadados, estando implementada a sua codificação em XML.

O padrão METS é projetado para funcionar como Pacotes de Informações OAIS, podendo assumir a função de SIP, AIP ou DIP dependendo da configuração das metainformações que serão inseridas e trabalhadas no escopo das seções do METS (LAVOIE, GARTNER, 2013; SARAMAGO, 2004).

A estrutura do METS, bem como a sua codificação em XML, é capaz de refletir as categorias de PDI exigidas ao AIP OAIS e permite a manutenção de presunção de autenticidade, confiabilidade e interoperabilidade a documentos arquivísticos.

Fluxos de trabalhos com os documentos arquivísticos videográficos da VSD, exigem um conjunto de distintas ferramentas de preservação digital para a execução de diferentes ações. Essas ferramentas e ações são capazes de gerar informações de PDI com padrões de metadados específicos para o contexto técnico dos vídeos. Essas metainformações devem ser registradas em padrões como o METS. (RICE, 2019)

#### *4.1 Padrões para documentos audiovisuais*

No campo da preservação de documentos audiovisuais existe uma vasta relação de padrões de metadados voltados para as mais diferentes especificidades desse gênero. Para além dos padrões/ esquemas de metadados discutidos anteriormente, apontaremos nessa seção alguns padrões que visam complementar possíveis lacunas de metainformação de preservação arquivística para documentos audiovisuais/ videográficos.

Como já mencionado, documentos videográficos apresentam uma gama de especificidades, sobretudo técnicas, que muitos dos padrões mais usuais de metadados não abarcam em seus elementos. Nesse sentido, os padrões de



metadados para documentos audiovisuais aqui discutidos, podem e devem ser entendidos como um possível componente agregador de metainformação aos demais padrões já consolidados no campo da preservação digital.

O padrão PBCore (Public Broadcasting Metadata Dictionary) foi criado pela comunidade de radiodifusão pública nos Estados Unidos da América para uso por emissoras públicas e outros. PBCore foi projetado para permitir a descrição de conteúdos audiovisuais produzidos por emissoras públicas no EUA.

O padrão abarca uma gama de metadados técnicos relativamente detalhados assim como metadados descritivos e de direitos. Muitos arquivos e bibliotecas tem adotado o uso do padrão que tem como base para o desenvolvimento do seu set de elementos o esquema Dublin Core.

Outro padrão consolidado para o registro de informações descritivas e técnicas de documentos audiovisuais é o padrão mantido pela European Broadcasting Union (EBU), designado EBUCore. Esse padrão é desenvolvido para descrição de mídia de transmissão de materiais audiovisuais na Europa.

No ano de 2015, as comunidades PBCore e EBUCore formaram um grupo de trabalho para melhorar as práticas de interoperabilidade entre os dois padrões. Como resultado, a comunidade PBCore decidiu pela adoção da ontologia do EBUCore e unificando os vocabulários controlados de ambos os padrões. O Manual de Catalogação de Imagem em Movimento da FIAF reconhece e incentiva o uso dos padrões PBCore e EBUCore.

O padrão AES57-2011 também é mencionado em boa parte da literatura pesquisada. Esse padrão é um padrão desenvolvido pela Audio Engineering Society (AES) com o objetivo de fornecer informações com alto nível de detalhamento sobre materiais de áudio.

Outro padrão associado a conteúdos audiovisuais é o padrão MPEG-7. É um padrão internacional chamado “Interface de descrição de conteúdo multimídia” (ISO/IEC 15938). Disponibiliza um conjunto de ferramentas de descrição audiovisual, descritores e esquemas de descrição para um acesso eficaz e eficiente (pesquisa, filtragem e navegação) a conteúdos multimídia.

É um padrão capaz de dar suporte ao desenvolvimento de banco de imagens para documentos audiovisuais. Entretanto, devido a sua complexidade não é muito recomendado para instituições, organizações ou pessoas custodiadoras de acervos audiovisuais que não tenham recursos suficientes para implementação e aproveitamento do uso de MPEG-7.

Um padrão pouco conhecido, mas não menos importante e indicado pela literatura pesquisada é o reVTMD (Reformatted Video Technical Metadata Standard). O padrão foi criado pela Audiovisual Preservation Solutions (AVP) em nome do National Archives and Records Administration (NARA, EUA) e em colaboração com o Grupo de Trabalho FADGI AV. Ele foi projetado especificamente para atender às



necessidades de objetos de vídeo reformatados dentro da comunidade do patrimônio cultural.

O padrão se apresenta em forma de esquema e não possui uma documentação estabelecida. Geralmente é utilizado como uma extensão de outros padrões como PBCore e PREMIS em METS para histórico de processos de reformatação com o objetivo de garantir que os representantes digitais gerados sejam autênticos. Essa problemática vai ao encontro do contexto observado para a digitalização da massa documental videográfica registrada em fitas magnéticas do acervo VSD.

## 6 CONCLUSÃO

Focando especificamente nos documentos audiovisuais videográficos característicos do acervo da VSD frente à literatura mais tradicional e a mais atualizada no campo da preservação audiovisual, o artigo busca propor a formulação de um perfil de padrões de metadados que atenda aos requisitos de preservação digital pautados em princípios arquivísticos.

A evolução tecnológica que impacta sobremaneira a produção, o gerenciamento e a preservação de coleções audiovisuais em meio digital, ratifica a necessidade da aplicação de ações práticas cada vez mais equilibradas entre a teoria arquivística tradicional e os novos paradigmas gerados por cenários cada vez mais impactados pelo uso de novas tecnologias de comunicação e informação. Nesse contexto, as estratégias de preservação de acervos documentais audiovisuais estão sendo constantemente remodeladas e atualizadas.

Especificamente para os documentos videográficos digitais da VSD foram indicados padrões fortemente reconhecidos e documentados no âmbito da descrição arquivística, sendo recomendados no escopo do perfil proposto os padrões *Dublin Core* complementado por elementos do EAD. Essa complementação do DC com elementos de EAD se deve ao fato do primeiro não abarcar elementos de descrição arquivística suficientes e baseados em normas como a Isad-G e a Nobrade. A complementação também observou a não repetitividade de metainformação em elementos distintos de ambos os padrões.

Para os metadados estruturais, foi indicado o padrão METS em função da capacidade de abarcar em suas 7 (sete) seções, os mais variados padrões de metadados em função das tipologias - descritivos, administrativos e estruturais - e finalidades de uso dos mesmos. Por essa razão, o METS é entendido e pode funcionar com um pacote (OAIS) SIP, AIP e DIP (LOC, METS: Introdução & Tutorial, online).

Para os metadados administrativos (técnicos, preservação e direitos) foram definidos os padrões PBCore e PREMIS. O padrão PBCore foi indicado por abarcar

melhor as especificidades para o registro de elementos (*instantiationEssenceTrack*<sup>5</sup>) de metadados técnicos de documentos videográficos como os da VSD.

Para os metadados de preservação o padrão PREMIS foi indicado por ser uma especificação abrangente para implementação de metadados de preservação digital que é revisada com base na experiência concreta e nas mudanças constantes dos ambientes tecnológicos de preservação.

Em resumo, o perfil de padrões de metadados que melhor atendeu a requisitos de preservação digital para os documentos videográficos da VSD (estruturados em pacotes AIP OAIS) com base na literatura e nos testes práticos da pesquisa está descrito a seguir:

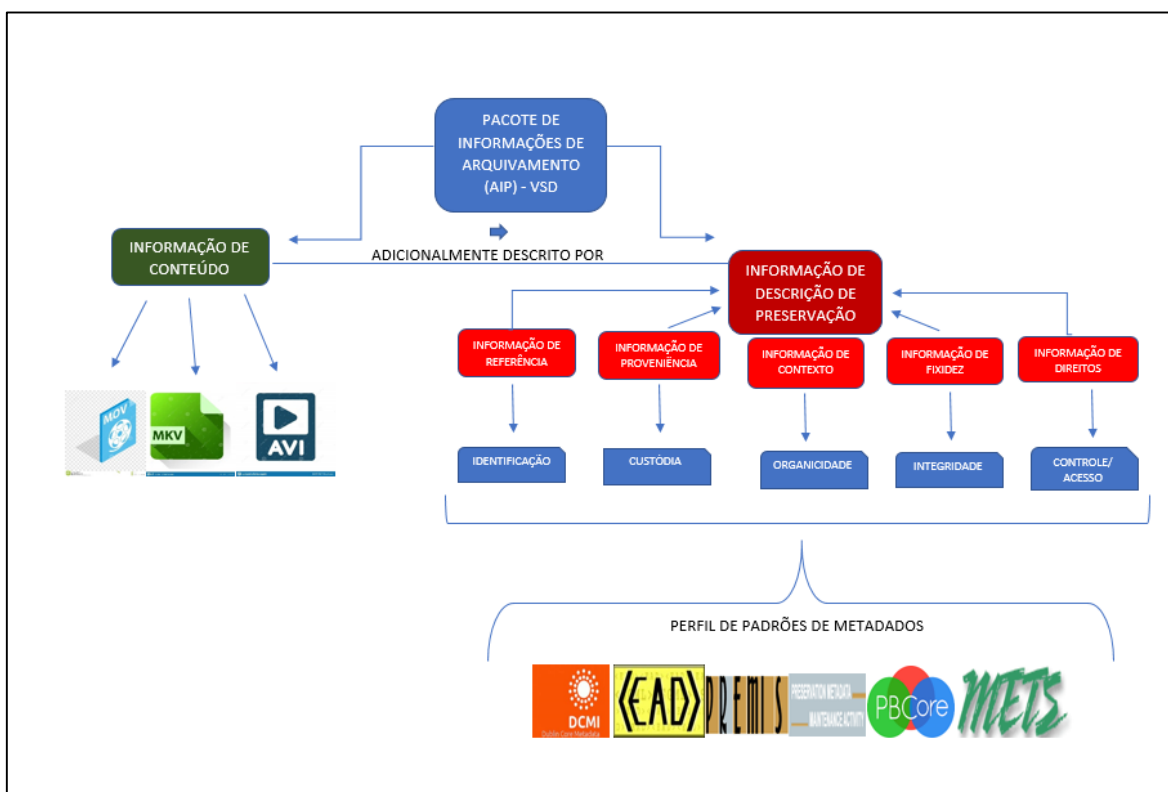
1. Metadados descritivos: DUBLIN CORE + EAD (Encoded Archival Description)
2. Metadados estruturais: METS (Metadata Encoding & Transmission Standard)
3. Metadados administrativos: PBCore (Public Broadcasting Metadata Dictionary) e PREMIS (Preservation Metadata: Implementation Strategies)

A figura 1 a seguir, demonstra um pacote AIP e a aplicação do perfil de padrões de metadados na estrutura das categorias de informação de PDI exigidas para garantir a confiabilidade e presunção de autenticidade dos documentos videográficos da VSD em conformidade com o OAIS:

---

<sup>5</sup> *instantiationEssenceTrack* é um elemento de contêiner XML que permite o agrupamento de elementos *essentialTrack* relacionados e seu uso repetido. Use o elemento *instantiationEssenceTrack* para descrever os fluxos individuais que compõem uma instanciação, como áudio, vídeo, código de tempo etc. Consultar em : <https://pbcore.org/elements/instantiationessencetrack.html>.

Figura 1: Perfil de padrões de metadados aplicados às PDI de um pacote AIP OAIS da VSD



**Elaboração:** João Guilherme Nogueira Machado (2022).

Longe de esgotar as discussões a respeito dos metadados no âmbito das demandas cada vez mais iminentes pelos acervos que compõem o patrimônio científico e cultural da Fiocruz, o artigo compreendeu bem a reflexão feita por GILLEAND (2007):

O que sabemos é que a existência de muitos tipos de metadados será fundamental para a continuidade da acessibilidade e utilidade *online* e intelectual dos recursos digitais e dos objetos de informação que eles contêm, bem como dos objetos e coleções originais aos quais se relacionam. Nesse sentido, os metadados nos fornecem a pedra de Roseta que permitirá decodificar objetos de informação e sua transformação em conhecimento nos sistemas de informação do patrimônio cultural do futuro. (GILLIAND, 2016, *online* apud BACA, 2007, tradução nossa)<sup>6</sup>

Nesse sentido, o entendimento de que estudos sobre a aplicabilidade de metadados - e as tecnologias que envolvem os seus usos - devam ser programados e atualizados enquanto estratégia fundamental para ações de preservação e acesso a longo prazo aos acervos científicos e culturais da instituição, além de propiciarem sustentabilidade efetiva e eficiente ao Programa de Preservação Digital da Fiocruz.

<sup>6</sup> What we do know is that the existence of many types of metadata will prove critical to the continued online and intellectual accessibility and utility of digital resources and the information objects that they contain as well as the original objects and collections to which they relate. In this sense, metadata provides us with the Rosetta stone that will make it possible to decode information objects and their transformation into knowledge in the cultural heritage information systems of the future.

## REFERÊNCIAS

BACA, Murtha. SHERMAN, Clarke. FRBR and Works of Art, Architecture, and Material Culture. *In*: TAYLOR, Arlene G. **Understanding FRBR**. Westport, CT: Libraries Unlimited, 2007. p.103-110. Disponível em: <https://bit.ly/42hBjSH>. Acesso em: 5 mai. 2023.

BUARQUE, Marco Dreer; MACHADO, João Guilherme Nogueira; PONTES, Eliane Batista. Plano de Preservação Digital da VideoSaúde: estratégias para a gestão de documentos audiovisuais. **RECIIS: Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 14, n. 3, p. 734-750, jul./set. 2020.

DOBREVA-MCPHERSON, Milena Petrova; KIM, Yunhyong; ROSS, Seamus. Metadata Extraction and Digital preservation: an overview. CONFERENCE ON DIGITAL LIBRARIES. Nice/França, 2007, 2, **Anais...** Disponível em: <https://bit.ly/3WWM6Rz3>. Acesso em: 5 mai. 2023.

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

FORMENTON, Danilo et al. Os padrões de metadados como recursos tecnológicos para a garantia da preservação digital. *Biblios [online]*, Pittsburgh, n.68, p.82-95, jul./set. 2017. Disponível em: <https://bit.ly/3WHz0Y5>. Acesso em: 5 mai. 2023.

FUNDAÇÃO OSWALDO CRUZ (FIOCRUZ). Casa de Oswaldo Cruz (COC). **Programa de Preservação Digital de Acervos da Fiocruz**. Rio de Janeiro: Fiocruz/COC, 2020. 37 p. Disponível em: <https://bit.ly/3WLUgMb>. Acesso em: 5 mai. 2023.

GRÁCIO, José Carlos Abbud. **Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior**. São Paulo, SP: Cultura Acadêmica, 2012. *E-book*. Disponível em: <https://bit.ly/3N4WqDB>. Acesso em: 05 maio 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 14721: 2012**. Space data and information transfer systems: open archival information system – Reference model. ISO: Genebra, 2012. Disponível em: <https://bit.ly/3qj0S8O>. Acesso em: 5 mai. 2023.

MÁRDERO ARELLANO, Miguel Ángel. **Critérios para a preservação digital da informação científica**. Brasília: UnB, 2008. 354f. Tese (Doutorado) - Programa de

Pós-Graduação em Ciência da Informação, Universidade de Brasília, Brasília, DF, 2008.

MÁRDERO ARELLANO. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, 2004. Disponível em: <https://bit.ly/3WZW8kV>. Acesso em: 5 mai. 2023.

RICE, Dave; SCHWEIKERT, Annie. Microserviços em Arquivos Audiovisuais. **Journal da Associação Internacional de Arquivos Sonoros e Audiovisuais (IASA)**, n.50, p.53–75, 2019.

SACRAMENTO, Carolina; BAIÃO, Fernanda de Araújo; FERREIRA, Simone Bacellar Leal. Um estudo sobre metadados para preservação e integração de acervos digitais heterogêneos. **RelaTe-DIA**, Rio de Janeiro, v.8, n.1, 2015. Disponível em: <https://bit.ly/3qj1tHA>. Acesso em: 5 mai. 2023.

SARAMAGO, Maria de Lurdes. Metadados para a preservação digital e aplicação do Modelo OAIS. *In*: CONGRESSO NACIONAL DE BIBLIOTECARIOS, ARQUIVISTAS E DOCUMENTALISTAS, Estoril, 2004, 8, **Anais...** Estoril: [s.n.], 2004, p.1-6. Disponível em: <https://bit.ly/3qbApKb>. Acesso em: 5 mai. 2023.

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://tinyurl.com/yk5a98kf>. Acesso em: 5 mai. 2023.