



XXIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – XXIV ENANCIB

ISSN 2177-3688

GT Especial

**PRINCÍPIOS DO OPEN GLAM E METADADOS EM COLEÇÕES RARAS E ESPECIAIS NA WEB:
UM ESTUDO SOBRE A BIBLIOTECA MÁRIO DE ANDRADE**

***OPEN GLAM PRINCIPLES AND METADATA IN RARE AND SPECIAL COLLECTIONS ON THE
WEB: A STUDY OF THE MÁRIO DE ANDRADE LIBRARY***

Gabriela Aparecida da Cunha Yamane – Universidade Federal de São Carlos (UFSCar)

Fabiano Ferreira de Castro – Universidade Federal de São Carlos (UFSCar)

Felipe Augusto Arakaki – Universidade Federal de São Carlos (UFSCar)

Felipe Ivo da Silva – Universidade Federal de São Carlos (UFSCar)

Rhuan Henrique Alves de Oliveira – Universidade Estadual Paulista (UNESP)

Modalidade: Trabalho Completo

Resumo: O problema de pesquisa aborda como as instituições de patrimônio cultural têm disponibilizado seu catálogo *online* e coleções de obras raras digitalizadas e disponibilizadas na *web*, quais padrões de metadados são utilizados e como ocorre o processo de escolha dos metadados com base em instrumentos de descrição para obras raras e especiais. O objetivo é discutir o compartilhamento de dados abertos em coleções de obras raras e especiais. Com base na revisão de literatura e em um estudo empírico na Biblioteca Mário de Andrade (BMA), foi possível identificar o processo de descrição dos metadados utilizados no catálogo *online* e nas coleções de obras raras digitalizadas que são disponibilizadas na *web*. Os resultados indicam que, embora a descrição no catálogo *online* esteja em consonância com o código de catalogação para obras raras, não foi possível identificar um padrão de metadados interoperável com outros sistemas, como ocorre na plataforma *online*. A análise dos dados demonstra a possibilidade de conflitos entre as orientações de segurança e os princípios do *OPEN GLAM*. Recomenda-se, que a BMA adote práticas mais consistentes para a gestão de metadados em seu catálogo *online*, bem como a implementação de políticas de catalogação e segurança para descrever obras raras e especiais conforme os padrões internacionais, como o DCRMb, alinhando-se aos princípios de acesso e de transparência do *OPEN GLAM*.

Palavras-chave: dados abertos; obras raras e especiais; metadados; catálogo *online*; biblioteca digital.

Abstract: The research problem addresses how cultural heritage institutions have made their online catalogues and collections of rare works digitized and available on the web, which metadata standards are used, and how the process of selecting metadata based on description tools for rare and special works takes place. The aim is to discuss the sharing of open data in collections of rare and special works. Based on a literature review and an empirical study at the Mário de Andrade Library (BMA), it was possible to identify the process of describing the metadata used in the online catalog and in the

collections of digitized rare works made available on the Web. The results show that although the description in the online catalog is in line with the cataloging code for rare works, it was not possible to identify a metadata standard that is interoperable with other systems, as is the case with the online platform. The analysis of the data shows the possibility of conflicts between the security guidelines and the principles of OPEN GLAM. It is recommended that the BMA adopt more consistent practices for managing metadata in its online catalog, as well as implementing cataloging and security policies to describe rare and special works according to international standards, such as DCRMB, in line with OPEN GLAM's principles of access and transparency.

Keywords: open data; rare and special materials; metadata; online catalog; digital library.

1 INTRODUÇÃO

O paradigma da tecnologia da informação, conforme Werthein (2000), destaca as transformações tecnológicas no âmbito social e o papel central da informação nas tecnologias contemporâneas, opondo-se às abordagens de períodos anteriores ao século XX, nas quais a tecnologia era adaptada para manipular a informação. Nas Bibliotecas Digitais (BD), por exemplo, essas mudanças se refletem na capacidade desses sistemas, enquanto ambientes informacionais, de facilitar a comunicação entre pesquisadores e as obras presentes em um acervo, de modo a intermediar esse processo.

As instituições responsáveis por patrimônio cultural¹ e bibliográfico e pela gestão de acervos raros e especiais mantêm informações primárias registradas em grande variedade de suportes. De acordo com Faria e Pericão (2008, p. 637), as coleções de livros raros ou reservas de livros raros são:

Onde estão depositados, em condições especiais, os incunábulos, as obras publicadas antes de 1800, as primeiras edições de textos científicos, literários e artísticos importantes, as obras com encadernações artísticas, de qualidade e de luxo e os exemplares únicos de uma edição esgotada, as obras autografadas ou as obras manuscritas que pertenceram a personagens célebres que por estas razões, pelo seu valor intrínseco e extrínseco, dimensão, encadernação, fragilidade ou por outros motivos justificam sua colocação em um lugar reservado (Faria; Pericão, 2008, p. 637).

¹ “O patrimônio cultural é composto por monumentos, conjuntos de construções e sítios arqueológicos, de fundamental importância para a memória, a identidade e a criatividade dos povos e a riqueza das culturas. Esta composição está definida na Convenção para a Proteção do Patrimônio Mundial, Cultural e Natural, elaborada na Conferência Geral da Organização das Nações Unidas para Educação, a Ciência e a Cultura (Unesco), em Paris (França), em 1972, e ratificada pelo Decreto No. 80.978, de 12 de dezembro de 1977”. Disponível em: <http://portal.lphan.gov.br/pagina/detalhes/29>. Acesso em: 20 jun. 2024.

Assim, justifica-se a importância de documentar a trajetória desses materiais dentro de uma instituição com fins patrimoniais. Nesse contexto, destaca-se a relevância da disponibilização, acesso, uso e reuso aos recursos do patrimônio cultural e bibliográfico por meio da digitalização e acesso em catálogos automatizados dessas coleções. A partir do século XX, de acordo com Borgan (1996), quando os *Online Public Access Catalogs* (OPACs) surgiram, ofereceram pouca melhoria em relação aos catálogos em fichas; porém, essa evolução possibilitou um acesso mais fácil às informações por pesquisadores em ambientes digitais com mais facilidade.

Desde o início do século XXI, os debates sobre Dados Abertos têm sido fundamentados no movimento de acesso aberto, ampliando sua importância, principalmente, com a declaração da *Budapest Open Access Initiative* (BOAI) em 2002. A esse respeito, tem-se que o Acesso Aberto (AA), de acordo com a BOAI (2002), pode ser definido como:

[...] literatura científica revisada por pares significa a disponibilidade livre na Internet, permitindo a qualquer usuário ler, fazer download, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral desses artigos, recolhe-los para indexação, introduzi-los como dados em software, ou usá-los para outro qualquer fim legal, sem barreiras financeiras, legais ou técnicas que não sejam inseparáveis ao próprio acesso a uma conexão à Internet. As únicas restrições de reprodução ou distribuição e o único papel para o direito autoral neste domínio é dar aos autores o controle sobre a integridade do seu trabalho e o direito de ser devidamente reconhecido e citado (BOAI, 2002).

Albagli, Clinio e Raychtock (2014, p. 440) sugerem que a expressão “dados abertos” no contexto governamental está associada à transparência dos dados, enquanto a Open Knowledge Foundation (OKF) estabelece seis categorias de dados abertos aplicáveis em diferentes áreas, como: “cultura, ciências, finanças, estatísticas, clima e meio ambiente” (OKF, [2024], tradução nossa). No âmbito cultural, esses dados se referem principalmente às descrições sobre materiais e artefatos culturais, que são administrados por instituições como galerias, bibliotecas, arquivos e museus.

Enquanto as discussões sobre acesso aberto defendido pela BOAI (2002) centram-se, primordialmente, na disponibilidade dos resultados de pesquisa, majoritariamente por meio de artigos publicados, os dados abertos visam a publicização de dados primários coletados durante a realização da pesquisa.

Assim, verifica-se que a iniciativa de acesso aberto, conforme a BOAI (2002), evidencia a importância da disponibilidade gratuita da literatura científica revisada por pares na internet, permitindo seu acesso, uso e reuso. No entanto, sob outra perspectiva, a (OKF, [2024]) amplia o escopo dos dados abertos para além do científico, incluindo outras categorias de dados, como os dados culturais.

Tal perspectiva suscita a necessidade de investigar como as instituições de patrimônio cultural têm disponibilizado seus catálogos no ambiente digital, com foco no escopo da Biblioteca Mário de Andrade² (BMA), especialmente para obras raras e especiais. Além disso, levanta questões sobre a escolha dos campos de descrição e os padrões de metadados adotados.

Dessa forma, o objetivo desta pesquisa consiste em discutir o compartilhamento de dados abertos em coleções de obras raras e especiais por meio da identificação e da descrição dos metadados utilizados em catálogos *online* e coleções de obras raras digitalizadas e disponibilizadas na *web*. O recorte investigado corresponde a cinco materiais raros da Coleção Brasileira, sendo que um deles, segundo Moraes *et al.* (1983, p. 743), configura-se como de extrema raridade, com data provável de 1625. Trata-se de um dos primeiros depoimentos que surgiram após a retomada da Bahia pelos espanhóis, que havia sido capturada pelos holandeses, descrevendo a chamada Jornada dos Vassalos. Além disso, a pesquisa também visa verificar se os metadados publicados comprometem a segurança e se as práticas do *Open Galleries, Libraries, Archives, and Museums* (OPEN GLAM) estão sendo atendidas. O *Open Glam*³ representa um movimento colaborativo entre indivíduos e organizações dedicadas à preservação do patrimônio cultural. Um dos principais objetivos é promover políticas de acesso aberto ao patrimônio cultural.

A relevância desta pesquisa integra o contexto crescente de digitalização e disponibilização de coleções de obras raras e especiais na *web*, bem como na necessidade de compreender como os metadados podem facilitar o compartilhamento desses materiais de forma aberta e acessível. A análise da utilização de metadados em catálogos *online* e coleções digitalizadas contribui para uma melhor compreensão das práticas contemporâneas de

² Site BMA: <https://www.prefeitura.sp.gov.br/cidade/secretarias/cultura/bma/>. Acesso em: 20 jun. 2024.

³ What Open Glam: <https://openglam.org/what/> acesso em: 14 set. 2024.

disseminação e preservação do patrimônio cultural na era digital, além de propor novas discussões.

Neste contexto, esta pesquisa justifica-se por sua contribuição ao campo acadêmico ao investigar como instituições de patrimônio cultural disponibilizam seus acervos *online*, discutindo padrões de metadados e práticas de descrição. No âmbito profissional, propõe melhorias na organização e no acesso aberto de coleções digitais. Socialmente, promove o acesso a materiais culturais, fomentando educação e pesquisa, e preservando a história e cultura para futuras gerações através do compartilhamento de dados abertos.

2 DADOS ABERTOS E SEGURANÇA DE ACERVO

A Biblioteconomia e a Ciência da Informação se estabeleceram para atender às necessidades de gestão e acesso à informação, impulsionadas pela explosão informacional, e investigam novas abordagens e ferramentas para lidar com o volume e a diversidade de dados. Estas investigações objetivam abordar questões fundamentais associadas ao acesso, obtenção, organização, disseminação e gerenciamento de registros de informação em uma variedade de suportes informacionais (Araújo, 2017; Saracevic, 2009).

Anteriormente ao ambiente *web*, quando os catálogos de bibliotecas não estavam disponíveis *online*, como é comum na contemporaneidade, Cutter (1876), conforme Svenonius (2000), fez a primeira declaração sobre os objetivos do sistema bibliográfico, baseando-se em três necessidades principais. O primeiro objetivo, de localização, auxiliava o usuário a encontrar um documento específico com base em informações como autor, título ou assunto. O segundo, de agrupamento, visava reunir todos os documentos de um autor, sobre um tema ou em um gênero. O terceiro, de escolha, auxiliava o usuário a encontrar entre documentos semelhantes, como diferentes edições de uma mesma obra. Em 1961, Lubetzky modificou esses objetivos durante a Conferência sobre Princípios de Catalogação em Paris, enfatizando a organização por autor e título. Já em 1997, a *International Federation of Library Associations and Institutions* (IFLA) propôs uma reformulação, considerando o contexto digital e a diversidade de materiais, com objetivos de localizar entidades (livros, artigos, etc.) que atendessem aos critérios de busca, identificá-las e distingui-las de outras semelhantes, selecionar a entidade mais adequada e adquiri-la ou acessá-la (Svenonius, 2000).

Os objetivos de Cutter (1876), Lubetzky (1961) e da IFLA (1997) destacam a importância histórica de registrar obras com detalhes de autor, título, assunto, conteúdo e edições disponíveis em instituições. Adequar esses objetivos para o contexto contemporâneo de dados abertos em coleções de obras raras e especiais (CORE) envolve questões suplementares, como a criação de metadados detalhados e acessíveis, os quais devem respeitar a singularidade das obras e a segurança dos acervos. No contexto da *web*, o processo progressivo de digitalização do patrimônio cultural e bibliográfico, tanto no âmbito nacional quanto internacional, tem instigado debates sobre o acesso aberto, disponibilização, localização, utilização e reutilização de documentos inseridos no ambiente digital.

Nesses ambientes digitais, os registros da história e da memória social além de desempenhar a função de preservação, também oferecem acesso e oportunidades de pesquisa em documentos únicos, ou seja, acesso a fontes primárias. Esse acesso permite o reconhecimento dos contextos culturais, históricos e sociais.

Nessa perspectiva, observa-se a ideia de biblioteca digital como um *locus* de interação entre pesquisadores e as fontes primárias documentais. No que concerne à sua origem, conforme Sayão (2009, p. 7):

As bibliotecas digitais surgem num contexto que sobrepõe, por um lado, a integração e uso das tecnologias de informação e de comunicação, das redes de computadores, das tecnologias de apresentação e o barateamento dos meios de armazenamento em massa; e, por outro, a disponibilidade crescente de conteúdos digitais em escala planetária, a possibilidade de digitalização a um custo economicamente viável de conteúdos em mídias convencionais e, *web* ainda, o fenômeno conhecido como coerência das mídias digitais, que abre a possibilidade singular para a concepção de novos serviços de informação a partir da integração de objetos digitais heterogêneos (Sayão, 2009, p. 7).

Além das questões de interação entre o acervo e os pesquisadores apresentadas por Sayão (2009), é necessário considerar a segurança desses acervos no que diz respeito à disponibilização dos metadados. Diante das discussões sobre acesso aberto e a disponibilização de metadados de materiais raros, a proteção dessas informações contra potenciais ameaças, como a preservação da integridade dos registros e a falsificação de documentos, bem como o tráfico ilícito de obras, destaca-se como preocupação primordial.

A IFLA (2023) destaca a descrição detalhada das coleções de obras raras e especiais como uma competência fundamental para os profissionais da área de Ciência da Informação.

As diretrizes propostas sugerem a criação de metadados detalhados, sendo estas essenciais para a identificação das coleções institucionais, considera ainda, que os requisitos devem ser específicos e variam conforme o formato dos itens. A descrição detalhada de características, como encadernação, marcas de propriedade e imperfeições é fundamental para garantir a preservação, a integridade e a história das obras, além de facilitar a distinção entre edições e impressões, por isso pode ser considerada uma medida e estratégia de segurança impreteríveis para prevenir e responder a vandalismos e furtos. Logo, segurança, preservação da autenticidade e da confiabilidade desses dados de acervos disponibilizados em ambientes *online* requer estratégias bem definidas de segurança, a fim de salvaguardar não apenas a sua existência, mas seu acesso para gerações futuras de pesquisadores.

No que concerne à segurança, esta pode ser compreendida como um conjunto de normativas adotadas para proteger os itens de uma coleção contra roubo, furto ou falsificação. Spinelli Junior (2009, p. 61) entende segurança de acervo como um grupo de “[...] elementos que formam um plano definido para impedir danos e combater os agentes prejudiciais à nossa instituição, abrangendo a proteção do edifício, dos bens, dos acervos e dos usuários”.

O *Rare Books and Manuscripts Section*⁴ (RBMS), um grupo da *American Library Association* (ALA) responsável por livros raros e coleções especiais, possui comitês que propõem diretrizes para profissionais e instituições. Em 2009, o *Security Committee* aprovou diretrizes revisadas em 2023, visando à preservação e o acesso contínuo às coleções especiais.

A *Association of College and Research Libraries* - ACRL (2023) na seção 6, intitulada “*Collections Workflows*”, discute questões de Catalogação. De acordo com as diretrizes, a representação e a descrição das coleções de uma instituição são consideradas fundamentais para a retenção e segurança. Nesta seção, o comitê afirma que os profissionais dessas coleções devem ser capazes de identificar os materiais por meio da catalogação e descrição, a fim de estabelecer perdas e fundamentar reivindicações para a recuperação de bens furtados, por exemplo.

No entanto, o comitê também ressalta que nem todos os dados de um registro devem ser disponibilizados ao público no catálogo *online*. Os dados que contêm informações específicas sobre uma cópia, por exemplo, usados para identificação em situações de perda

⁴ Site RBMS: <https://rbms.info/> Acesso em: 20 jun. 2024.

ou roubo, devem ser mantidos em uma área segura e separados de outros registros do catálogo ou arquivos curatoriais.

Nesse sentido, diante da relevância, utilidade e necessidades proporcionadas pelo acesso dos registros em coleções Raras e Especiais em ambientes digitais, as instituições que salvaguardam essa tipologia de acervos precisam discutir e propor constantemente diretrizes para o compartilhamento de seus metadados em seus catálogos *online* no contexto de dados abertos.

Em consonância com as discussões sobre dados abertos em acervos compostos por patrimônio cultural e bibliográfico, apresenta-se o movimento de iniciativa global denominado *OPEN GLAM*, cujo acrônimo representa: *Galleries, Libraries, Archives and Museums*. O movimento está fundamentado na ideia de que essas instituições desempenham um papel primordial no progresso da sociedade, devido aos registros históricos que salvaguardam e que constituem importantes fontes de estudo, de pesquisa e de memória.

À medida que as instituições responsáveis pelo patrimônio cultural, como as coleções de obras raras e especiais, tornam seus metadados e acervos digitalizados acessíveis *online*, oferecem à sociedade a compreensão da trajetória humana, além de incentivarem os pesquisadores a se engajarem ativamente, contribuindo, participando e compartilhando os dados advindos desse processo (Creative Commons Brasil, 2019).

Os princípios do *OPEN GLAM* orientam instituições culturais a adotar medidas para disponibilizar seus acervos de maneira acessível, permitindo ampla utilização por terceiros, incluindo traduções, remixagens e recriações. De acordo com Creative Commons Brasil (2019), as instituições que aderem à iniciativa *OPEN GLAM* defendem os seguintes princípios:

1. Disponibilizar informação digital sobre os objetos (metadados) no domínio público utilizando uma ferramenta legal apropriada como o instrumento do Creative Commons Zero. Isso promove a maior reutilização possível do dado e facilita a descoberta de seus recursos, ao mesmo tempo em que assegura a conformidade com os principais agregadores culturais, como a Europeana e a Digital Public;
2. Não adicionar novos direitos a representações digitais de obras cujos direitos autorais já tenham expirado (e que já estejam, portanto, em domínio público) de forma a mantê-las em domínio público. As cópias digitais e as representações de obras cujos direitos autorais já expiraram (obras que estão no domínio público) devem ser explicitamente assinaladas usando uma ferramenta legal apropriada como a marca de Domínio Público do Creative Commons;

3. Quando for publicar dados, fazer uma declaração explícita e robusta de suas vontades e expectativas com relação à reutilização e redirecionamento das descrições, do conjunto de dados como um todo e de subconjuntos do acervo;
4. Quando publicar dados usar formatos de arquivo abertos que sejam legíveis por máquina. Os formatos que são legíveis por máquinas são aqueles que são capazes de ter os seus dados extraídos por programas de computador. Se a informação é divulgada num formato de arquivo fechado, isso pode causar obstáculos significativos à reutilização da informação nele codificada, forçando aqueles que desejam utilizar a informação a comprar o software necessário. A estrutura e os possíveis usos dos dados devem ser bem documentados, por exemplo em um datablog ou em páginas *web*;
5. Oportunidades de envolver o público de novas formas na internet devem ser exploradas. Documente claramente os dados, conteúdos e serviços abertos que você fornece de forma a que outros possam facilmente reutilizar, aproveitar e melhorar o que você disponibilizou. Ao publicar dados, esteja disposto a responder perguntas de partes interessadas sobre o dado e a apoiá-las em aproveitar ao máximo seus dados. Ofereça oportunidades para que seu público possa fazer a curadoria e a coleção de itens de seus acervos. O Rijksstudio do Rijksmuseum é um ótimo exemplo desse tipo de colaboração. Quando possível, considere permitir que seus usuários enriqueçam e melhorem seus metadados utilizando aplicativos de crowdsourcing (Creative Commons Brasil, 2019).

Os princípios declarados no *OPEN GLAM*, encontram correlações com algumas questões sobre a segurança e a integridade dos acervos em ambientes digitais, como discutido anteriormente. Essas diretrizes convergem com questões-chave enfrentadas pelos profissionais de preservação patrimonial, principalmente em dois pontos: 1) Na necessidade de explicitar declarações sobre reutilização e redistribuição de informações do acervo, visando garantir segurança, confiabilidade e uso responsável e transparente dos dados; 2) Em relação aos novos direitos e domínio público, é essencial preservar a autenticidade e a integridade dos registros digitais, garantindo acesso, uso e reuso.

A convergência entre as diretrizes do *OPEN GLAM* e os desafios enfrentados na preservação de acervos raros e especiais ressalta a importância de estratégias que garantam tanto o acesso quanto a proteção a longo prazo desses recursos para as futuras gerações de pesquisadores.

De acordo com as características definidas por Faria e Pericão (2008, p. 637), aspectos como encadernação, fragilidade, especificidades sobre edições e marcas de proveniência são essenciais para comprovar a trajetória de uma obra dentro de uma instituição. Conforme Pinheiro, Helde e Pereira (2023, p. 177), essas marcas de proveniência podem se apresentar

“[...] sob a forma de assinaturas, autógrafos, carimbos, ex-donos, ex-líbris, super libros, etiquetas, dedicatórias, entre outros”.

A ACRL (2011) no *Descriptive cataloging of rare materials, books* (DCRMb) na regra “7B19.2.” orienta a descrição dos detalhes sobre a procedência dos itens em uma nota local, caso o catalogador julgue importante. Em descrições menos detalhadas, o código instrui a resumir alguns dados de procedência sem fornecer transcrições ou descrições exatas das evidências. Por exemplo, o “*Bibliographic Standards Committee*” da *Rare Books and Manuscripts Sections* (RBMS) apresenta materiais descritos no Formato MARC 21, onde é possível observar que algumas marcas, como as de propriedade, são descritas no campo 561##\$a, enquanto as descrições de encadernações são registradas em 563##\$a e algumas anotações são registradas em notas gerais (5XX). No entanto, omitir tais dados pode conflitar com os princípios do *OPEN GLAM* e gerar problemas de segurança de acervo, como abordado anteriormente.

Nesse âmbito, os metadados desempenham um papel essencial, pois podem proporcionar a gestão eficaz e a preservação a longo prazo dos objetos de acervos raros, possibilitando o acesso e a compreensão desses recursos informacionais.

2.1 Fundamentação dos metadados

De acordo com Alves (2010, p. 14), os metadados podem ser considerados “[...] atributos que representam um recurso informacional” com o objetivo de localizar e recuperar dados. Um metadado pode indicar um atributo que representa uma entidade, fornecendo dados compreensíveis por ferramentas computacionais. Em registros de obras raras, os metadados apresentam dados úteis como proveniência, origem, tipo de aquisição, encadernação e detalhes físicos específicos. Conforme apresentado por Alves (2010, p. 14), na descrição da informação, considera-se que:

Metadados se constituem como um elemento fundamental no processo de tratamento descritivo da informação, pois refletem a conjunção entre aspectos tecnológicos e de representação necessários aos novos tipos de recursos e ambientes informacionais surgidos, facilitando assim, uma recuperação mais eficiente em ambientes digitais (Alves, 2010, p. 14).

No contexto da definição do termo “metadados”, é importante considerar os padrões de metadados adotados e as funções que os metadados desempenham. Segundo Grácio

(2002, p. 21) metadados são conjuntos de elementos que descrevem o conteúdo dos recursos, permitindo que pesquisadores ou mecanismos de busca acessem e recuperem esses recursos. Assim sendo, a seleção de metadados para descrição é imperativa, visto que define o nível de detalhamento do material. A escolha desses elementos determina o nível de detalhamento da descrição, influenciando diretamente o grau de especificidade do registro. Nessa perspectiva, os dados e os metadados em obras raras são essenciais para a preservação do patrimônio cultural, científico e artístico, garantindo a continuidade e o enriquecimento do legado cultural e científico, quando adequadamente bem descritos e padronizados.

Há vários padrões de metadados disponíveis, cada um adequado para diferentes domínios e informações específicas. Nas bibliotecas digitais, o *Dublin Core* (DC) é amplamente adotado para promover a interoperabilidade entre sistemas, sendo aplicável a diferentes recursos, como documentos, imagens e objetos digitais. Ribeiro (2021, p. 73) destaca a necessidade de definir um padrão de metadados adaptado para obras raras. De acordo com o autor:

A definição e adaptação de um padrão de metadados para obras raras possibilitará o desenvolvimento de uma ferramenta de consulta e estudos, auxiliando na gestão desses acervos e na recuperação de informações. Identificar quais metadados serão necessários para descrever as obras raras na *web* será essencial. Para o ambiente da *web*, o padrão de metadados *Dublin Core* é hoje uma alternativa segura, interoperável e consolidada para a descrição de documentos analógicos e digitais (Ribeiro, 2021, p. 73).

O *Dublin Core* (DC) permite a escolha livre e repetida de elementos, sem uma ordem específica. Essa flexibilidade facilita a incorporação de elementos essenciais e adicionais, adequando-se às necessidades específicas da descrição e representação de obras raras e especiais *na web*.

3 PROCEDIMENTOS METODOLÓGICOS

Este estudo, caracterizado como uma pesquisa aplicada, adota uma metodologia exploratória e descritiva. Foram analisados os metadados de cinco registros presentes no catálogo *online*, bem como cinco objetos digitalizados e disponibilizados na plataforma *online*. Esses registros e objetos digitais pertencem ao acervo da Seção de Obras Raras e Especiais da Biblioteca Mário de Andrade, datam do período histórico de 1625, e fazem parte da coleção

Brasileira digitalizada pela instituição⁵. A seleção destes registros e objetos digitais visou contemplar a representatividade da coleção em análise no que concerne à coleção de obras raras e especiais (CORE).

A pesquisa foi estruturada em três etapas principais, as quais são especificadas a seguir: 1) A etapa de coleta de dados envolveu a extração e a documentação dos metadados associados aos registros selecionados no catálogo *online*. De maneira semelhante, foram coletados os metadados disponíveis na plataforma *online* DocPro dos objetos digitalizados no projeto “Tesouros da Cidade de São Paulo”. O ambiente foi escolhido porque permite a comparação de metadados entre objetos digitais e registros do OPAC, que estão disponibilizados por meio do sistema Alexandria; 2) Os metadados coletados foram, então, submetidos a uma comparação fundamentada nos padrões internacionais de descrição e representação de obras raras, como DCRMb e MARC 21. A comparação também incluiu a verificação da adequação dos metadados às práticas recomendadas pelo movimento de Dados Abertos e pela iniciativa GLAM. Em seguida, foi realizada uma comparação entre os metadados dos registros no catálogo *online* e os dos objetos digitalizados; 3) Por fim, com base nos resultados da análise, foram discutidas as práticas atuais de descrição e representação de obras raras e especiais. A discussão incluiu a análise das práticas no contexto dos Dados Abertos e do movimento GLAM, propondo recomendações para aprimorar a disponibilização dos metadados. As recomendações propõem uma reflexão sobre o acesso aberto aos metadados de coleções digitalizadas de obras raras e especiais, considerando que a disponibilização dessas informações pode comprometer a integridade dos objetos físicos.

4 RESULTADOS

A Biblioteca Mário de Andrade (BMA), fundada em 1925 como a Biblioteca Municipal de São Paulo, é uma instituição pública subordinada à Secretaria Municipal de Cultura da Prefeitura de São Paulo. Reconhecida como patrimônio cultural, é protegida pelas autoridades municipais de preservação histórica, conforme a Resolução SC-82, de 20 de agosto de 2013. Desde 1945, oferece acesso ao público à sua Seção de Obras Raras e Especiais, idealizada por Rubens Borba de Moraes. A Seção possui um diversificado acervo de livros, periódicos,

⁵ História da Biblioteca Mário de Andrade: <https://www.prefeitura.sp.gov.br/cidade/secretarias/cultura/bma/historico/index.php?p=7653> Acesso em: 20 jun. 2024.

cartões-postais, manuscritos, fotografias e outros materiais acessíveis por meio do catálogo *online* com o *software* Alexandria⁶.

Além de disponibilizar os metadados no OPAC, foram realizados três projetos de digitalização e disponibilização *online* do acervo. A instituição destaca essas iniciativas como principais por meio de projetos e parcerias conhecidos como: 1) “Tesouros da Cidade de São Paulo”⁷, que teve início em 2001, por meio da Secretaria Municipal de Cultura de São Paulo (SMC). Esse projeto propunha a criação de uma biblioteca virtual composta pelas coleções especiais da SMC. Foram envolvidos a SMC, o Instituto Embratel 21 e a plataforma DocPro, resultando na digitalização e na disponibilização *online* das coleções Brasileira (sobre o Brasil, 1562-1885), Fotos e Ilustrações de São Paulo (1568-1960) e Paulistanas (sobre a cidade de São Paulo, 1736-1934); 2) Em 2021, houve uma parceria com o Google Arts & Culture⁸, em que foram digitalizadas 639 obras, incluindo mapas, cartazes e fotografias; 3) Além dessas, houve uma parceria com a Biblioteca Digital UNESP⁹, na qual foi convertida uma amostra de periódicos da diversificada imprensa paulistana do início do século XX, do formato de microfilme para o digital.

No contexto do OPAC, o *software* Alexandria é um sistema gerenciador de bibliotecas desenvolvido em linguagem Java, disponibilizado na *web* como cliente multiplataforma, criado e atualizado pela empresa Docs & Bytes Informática Ltda. Para materiais raros e especiais, a BMA realiza a descrição, com entrada dos dados no formato MARC 21 através do *software* Alexandria, na descrição são utilizados principalmente os campos (MARC21 tag 563##; DCRM(B): 7B19.3.1) para as descrições de encadernações, e as notas gerais (MARC21 tag 500##; DCRM(B): 7B10.1) para outras características intrínsecas dos materiais. As marcas de propriedade e proveniência não são disponibilizadas aos pesquisadores no OPAC; um dos formatos de saída dos dados é realizado por um recurso de conversão para metadados com a representação no formato *Dublin Core*. A estrutura utilizada inclui o *Hyper Text Markup Language* (HTML) com o protocolo *Open Graph* (OG).

⁶ Catálogo Eletrônico Obras Raras e Periódicos: <http://bibliotecacircula.prefeitura.sp.gov.br/PesquisaRarosePeriodicos/> Acesso em: 20 jun. 2024.

⁷ Tesouros da cidade de São Paulo: <https://docvirt.com/demo/bma2/bma.htm>. Acesso em: 20 jun. 2024.

⁸ Google Arts & Culture: <https://artsandculture.google.com/partner/biblioteca-m%C3%A1rio-de-andrade> Acesso em: 20 jun. 2024.

⁹ Periódicos Paulistas: <https://bibdig.biblioteca.unesp.br/communities/4de7d0e7-fcea-4e57-b6d9-08749ee0d1b8?cmcl.page=1>.

O protocolo OG¹⁰ consiste em uma estrutura que possibilita a representação de um objeto gráfico por meio de metadados incorporados em uma estrutura HTML. Sua versão inicial é baseada em *Resource Description Framework in Attributes* RDFa, a qual permite a inserção de *tags* <meta> adicionais no <head> do código. Os metadados essenciais, os quais são obrigatórios, incluem: “og:title”; “og:type”; “og:image”; “og:url”.

No “Tesouros da cidade de São Paulo”, os materiais digitalizados foram disponibilizados na *web* apenas para visualização, Figura 1, sendo vedada a cópia por *download*, a partir do navegador por meio de programação.

Figura 1 – Acervo digital Biblioteca Mário de Andrade na Plataforma DocPro.



Fonte: Extraído de Tesouros da cidade de São Paulo (2024).

O sistema permite a pesquisa nos textos das imagens geradas a partir de papel e documentos digitais. A definição dos valores de descrição dos livros foi realizada de acordo com o MARC21 e o código DCRMb. Não foi possível identificar o uso de um padrão de metadados, como o *Dublin Core*, no código-fonte da plataforma DocPro. Observou-se a construção do *script* para entrada de dados no HTML com DocType; algumas dessas linhas de metadados podem ser identificadas na Figura 2, enquanto a descrição dos livros está incorporada em uma única tag HTML denominada "textarea name", conforme representado na Figura 3.

¹⁰ Protocolo *Open Graph*: <https://ogp.me/> Acesso em: 20 jun. 2024.

Figura 2 – Entrada de dados HTML DocType na Plataforma DocPro.

```
<title> Brasiliana - Coleção de livros sobre o Brasil - DocReader Web </title>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
<meta http-equiv="imagetoolbar" content="no">
<meta http-equiv="X-UA-Compatible" content="IE=edge,chrome=1">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
```

Fonte: Extraído de Tesouros da cidade de São Paulo (2024).

Figura 3 – Tag HTML “textarea name” na Plataforma DocPro.

```
<textarea name="TextoDigitadoTxt" rows="2" cols="20" readonly="readonly" id="TextoDigitadoTxt" style="height:110px;">
"Autor: RODRIGUEZ DE BURGOS, Bartholomé, séc.16-17. Título: Relacion de la jornada del Brasil, escrita a Juan de Castro / por Bartholome Rodriguez de Burgos. Local/Data: [Cadiz: por Juan de Borja, 1625] [4] p. Notas: Obra da maior raridade, não descrita por Palau; impressão retirada do colofão. Faz o relato da retomada da Bahia pelos espanhóis, após a captura pelos holandeses. Fontes: Maggs, v.5, p.203; RBM, v.2, p.743: "extremely rare". Bosch, p.92." == $0
</textarea>
```

Fonte: Extraído de Tesouros da cidade de São Paulo (2024).

A plataforma disponibiliza acesso aos objetos digitais, porém não é possível identificar um padrão que seja interoperável com outros sistemas. Ela apresenta uma funcionalidade de busca que utiliza *Optical Character Recognition* (OCR), o que pode auxiliar na pesquisa de documentos manuscritos e tipográficos, por exemplo.

5 CONSIDERAÇÕES FINAIS

O objetivo principal da pesquisa consistiu em discutir o compartilhamento de dados abertos e nos princípios do movimento *OPEN GLAM* em coleções de obras raras e especiais (CORE). Este objetivo foi alcançado por meio da identificação dos metadados utilizados no OPAC e na plataforma *online* de obras raras que foram digitalizadas pela Biblioteca Mário de Andrade (BMA) e disponibilizadas na *web*, especificamente aqueles que fazem parte da coleção Brasiliana, ano de 1625.

Em observância às questões apresentadas sobre a escolha dos campos, o DCRMb orienta a inclusão de detalhes sobre a origem dos itens na área de notas, possibilitando aos catalogadores o registro de informações como marcas de propriedade, tais como anotações manuscritas (MARC21 tag 561##; DCRM(B): 7B19.2), encadernações e outras características físicas extrínsecas (MARC21 tag 563##; DCRM(B): 7B19.3.1) e de origem ou intrínsecas que

sejam pertinentes (MARC21 tag 500##; DCRM(B): 7B10.1). O comitê de segurança de acervo da ACRL, por sua vez, considera que a representação e a descrição são fundamentais para a retenção e a segurança. No entanto, o mesmo comitê debate sobre a guarda de arquivos curatoriais para identificação de perda ou roubo. Neste ponto, conforme observado, existe uma divergência, pois a descrição menos detalhada dos elementos pode ser um ponto de conflito com os princípios do *OPEN GLAM*. O conflito pode surgir porque o comitê de segurança propõe a guarda curatorial das descrições, o que resultaria em uma representação menos detalhada para o pesquisador, visando proteger informações sensíveis que poderiam facilitar roubos ou perdas. Esta prática, no entanto, pode ir contra o primeiro princípio do *OPEN GLAM* de promover a maior reutilização possível dos dados e facilitar a descoberta dos recursos culturais. O *OPEN GLAM* enfatiza a importância de publicar metadados completos e detalhados que permitam uma maior reutilização e descoberta dos recursos culturais. Outro princípio que conflita é o terceiro, que defende uma declaração explícita e robusta das vontades e expectativas em relação à reutilização e redirecionamento das descrições e dos conjuntos de dados. Este sugere transparência e clareza nas intenções de reutilização dos dados, o que demanda descrições completas e acessíveis. Este movimento global requer transparência na descrição e disponibilização de metadados.

A utilização do formato MARC21 exemplifica como a Seção de Obras Raras e Especiais da BMA descreve e representa os materiais. São utilizados principalmente os campos para descrições de encadernações e características intrínsecas. No entanto, as marcas de propriedade e proveniência não são disponibilizadas no OPAC. Um dos princípios do movimento *OPEN GLAM* propõe disponibilizar informação digital sobre os objetos (metadados) no domínio público. No entanto, o compartilhamento de informações detalhadas sobre o acervo pode comprometer a segurança ao expor dados sensíveis ou que possam ser explorados indevidamente. Portanto, a transparência, aliada a medidas robustas de segurança são necessárias para proteger a integridade dos dados e preservar a herança cultural. Acredita-se que, para isso, seja importante que a biblioteca publique uma declaração explicando a reutilização dos metadados por meio de políticas, o que pode incluir instruções sobre como citar as fontes, como solicitar permissão para usos do acervo, entre outros casos.

Recomenda-se, que a BMA adote práticas mais consistentes para a gestão de metadados em seus ambientes digitais. No OPAC, é aconselhável que os metadados dos registros bibliográficos sejam disponibilizados com a indicação de *Licenças Creative Commons*,

priorizando o uso e a reutilização. Além disso, para obras cujos direitos autorais tenham expirado e estejam no domínio público, é importante que sejam explicitamente identificadas, utilizando a marca de Domínio Público do *Creative Commons* (CC0). Nas plataformas de objetos digitais, recomenda-se que a biblioteca adote formatos abertos e interoperáveis, garantindo que as descrições e as representações digitais das obras em domínio público estejam acessíveis e marcadas de forma adequada. É também aconselhável, que sejam feitas declarações explícitas de reutilização dos dados, tanto no catálogo *online* quanto nas plataformas de objetos digitais, promovendo transparência, acesso, uso e reuso dos metadados.

É fundamental implementar políticas de catalogação e segurança para descrever obras raras e especiais conforme os padrões internacionais como o DCRMb, alinhando-se aos princípios de acesso e de transparência do *OPEN GLAM*. Portanto, conclui-se, que as instituições devem adotar uma abordagem curatorial na qual devem garantir que os metadados fornecidos sejam suficientemente detalhados para fins educacionais e de pesquisa, ao mesmo tempo em que devem proteger a privacidade e a segurança das coleções.

REFERÊNCIAS

ALBAGLI, S. *et al.* Ciência aberta: correntes interpretativas e tipos de ação. **Liinc em revista**, Rio de Janeiro, v. 10, n. 2, 2014.

ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

ARAÚJO, C. A. A. Teorias e tendências contemporâneas da ciência da informação. **Informação em Pauta**, Fortaleza, v. 2, n. 2, p. 9-34, dez. 2017.

ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES. Rare Books and Manuscripts Section. **ACRL/RBMS guidelines regarding the security of special collections materials**. Rev. June 2023. Chicago: American Library Association, 5 out. 2009. Disponível em: https://www.ala.org/acrl/standards/security_theft. Acesso em: 24 jun. 2024.

ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES. Rare Books and Manuscripts Section. Bibliographic Standards Committee. **Descriptive cataloging of rare materials (books)**. Washington: Library of Congress, 2011. Disponível em: https://rbms-info.translate.googleusercontent.com/translate/pt-BR&_x_tr_pto=wapp. Acesso em: 15 jun. 2024.

BORGMAN, C. L. Why are online catalogs still hard to use?. **Journal of the American society for information science**, Hoboken, v. 47, n. 7, p. 493-503, July 1996.

BUDAPEST OPEN ACCESS INITIATIVE. **Read the declaration**. Budapest: BOAI, 14 Feb. 2002. Disponível em: <https://www.budapestopenaccessinitiative.org/read/>. Acesso em: 20 jun. 2024.

CREATIVE COMMONS BRASIL. **Os 5 princípios do Open Glam**. 2019. Disponível em: <https://br.creativecommons.net/2019/09/24/os-5-principios-do-open-glam/>. Acesso em: 20 jun. 2024.

FARIA, M. I.; PERICÃO, M. da G. **Dicionário do livro**: livro eletrônico. São Paulo: EDUSP, 2008.

GRÁCIO, J. C. A. **Metadados para a descrição de recursos da internet**: o padrão Dublin Core, aplicação e a questão da interoperabilidade. 2002. 127 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **Diretrizes sobre as competências do profissional responsável por livros raros e coleções especiais**. Den Haag: IFLA, 2023.

MORAES, Rubens Borba de. **Bibliographia brasiliana**: rare books about Brazil, published from 1504 to 1900 and works by Brazilian authors of the Colonial period. Rev. and enlarged ed. Los Angeles [Estados Unidos]: UCLA Latin American Center Publications; Rio de Janeiro, RJ: Kosmos, 1983. 2v.

OPEN KNOWLEDGE FOUNDATION. **What is open?**. England: OKF, [2024]. Disponível em: <https://okfn.org/opendata/>. Acesso em: 14 set. 2024.

PINHEIRO, A. de S.; HELDE, R. R. Von; PEREIRA, S. F. **Glossário ilustrado de livros raros e acervos de memória**. Rio de Janeiro: Fundação Biblioteca Nacional, 2023.

RIBEIRO, J. R. *et al.* **Contribuições metodológicas para criação de bibliotecas digitais de livros raros**: aplicação na coleção Brasiliana da UFMG. 2021. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2021. Disponível em: <http://hdl.handle.net/1843/36609>. Acesso em: 24 jun. 2024.

SÃO PAULO (SP). Secretaria Municipal de Cultura. Tesouros da Cidade de São Paulo. São Paulo: Secretaria Municipal de Cultura, [2024]. Disponível em: <https://docvirt.com/demo/bma2/bma.htm>. Acesso em: 12 jun. 2024.

SARACEVIC, T. Information science. In: BATES, M. J; MAACK, M. N. (ed.). **Encyclopedia of Library and Information Science**. 3rd. ed. New York: Taylor & Francis, 2009. p. 2570-2586.

SAYÃO, L. F. Afinal, o que é biblioteca digital?. **Revista USP**, São Paulo, n. 80, p. 6-17, 2009.

SPINELLI JÚNIOR, J. **Guia de preservação e segurança da Biblioteca Nacional**. Dissertação (Mestrado Profissional em Bens Culturais e Projetos Sociais) – Fundação Getúlio Vargas, Rio de Janeiro, 2009.

SVENONIUS, E. **The intellectual foundation of information organization**. London: MIT Press, 2000.

WERTHEIN, J. A sociedade da informação e seus desafios. **Ciência da Informação**, Brasília, v. 29, n. 2, p. 71-77, maio-ago. 2000.