

Integração de Acervos do Patrimônio Cultural:

Bases Teóricas para a Interoperabilidade

Rodrigo Moreira Garcia — Doutorando, Programa de Pós-Graduação em Ciência da Informação, Escola de Comunicações e Artes, Universidade de São Paulo (USP), Brasil • ORCID: <https://orcid.org/0000-0003-2694-5988> • garcia.rodrigo@usp.br

Francisco Carlos Paletta — Livre-Docente, Escola de Comunicações e Artes, Departamento de Informação e Cultura, Universidade de São Paulo (USP), Brasil • ORCID: <https://orcid.org/0000-0002-4112-5198> • fcpaletta@usp.br

1. Introdução

A institucionalização dos acervos culturais — compreendida como a incorporação sistemática e curatorial de conjuntos documentais em bibliotecas, arquivos e museus — assume importância estratégica nas políticas de preservação da memória e na democratização do conhecimento. Ao reconhecer esses materiais como patrimônio cultural e histórico, a atuação das instituições de custódia não apenas assegura sua permanência, como também amplia seu valor científico, educativo e simbólico. No entanto, a diversidade tipológica dos objetos e a fragmentação das abordagens descritivas entre os domínios arquivístico, biblioteconômico e museológico têm gerado ambientes informacionais isolados, dificultando a descoberta contextualizada de documentos, a integração entre coleções e o desenvolvimento de novas leituras sobre a memória social (Vital & Brascher, 2017, pp. 202-207; Marcondes, 2019, pp. 45-52).

Esse cenário caracteriza um problema de escala cultural: quanto maior a heterogeneidade dos acervos, maior o custo cognitivo e técnico para interpretá-los e conectá-los (Marcondes, 2019). As infraestruturas digitais que sustentam esses repositórios frequentemente refletem barreiras institucionais históricas, com sistemas baseados em padrões autônomos, desenvolvidos para finalidades internas específicas. Como consequência, observa-se a formação dos chamados *silos de dados*, nos quais conjuntos ricos em informação permanecem inacessíveis fora de seu contexto institucional original, limitando o potencial de reutilização, reinterpretação e visibilidade pública dos acervos (O'Neill & Stapleton, 2022, pp. 891-903).

No contexto internacional, especialmente em iniciativas como *Europeana*¹, *Digital Public Library of America*² (DPLA), *WarSampo Portal*³, *Trove*⁴ e *DigitalNZ*⁵, observa-se um movimento progressivo de ruptura com os modelos fechados e de transição para uma lógica de *web* semântica, orientada à publicação e interligação de metadados abertos. O uso de tecnologias como o *Resource Description Framework*⁶ (RDF), identificadores persistentes⁷ (URIs), ontologias formais e vocabulários controlados interligados por princípios de *Linked Open Data*⁸ (LOD) tem permitido o desenvolvimento de grafos de conhecimento que interconectam acervos diversos, promovendo um ambiente de descoberta transversal e de inferência semântica entre objetos culturais (Doerr, 2011, pp. 75-92; Concordia et al., 2010, pp. 61-69).

Contudo, essa transição não é meramente técnica. Exige uma reconfiguração profunda nas políticas institucionais, nas competências profissionais e nos modelos de governança das instituições de

memória. O'Neill e Stapleton (2022, pp. 891-903) destacam que romper os *silos* não é apenas conectar dados, mas ultrapassar as divisões disciplinares que historicamente segmentaram as abordagens de organização da informação. Nesse sentido, a interoperabilidade não deve ser reduzida à compatibilidade técnica de formatos, mas compreendida como um conceito ampliado, que envolve a construção de consensos semânticos, a adoção de princípios FAIR⁹ (*Findable, Accessible, Interoperable, Reusable*) e o incentivo a práticas colaborativas e interdisciplinares.

No Brasil, embora existam experiências relevantes em iniciativas como a BBM Digital¹⁰, o Tainacan¹¹, a BNDigital¹² e a Brasileira Iconográfica¹³, ainda persistem desafios para a consolidação de uma política nacional de integração de acervos culturais baseada em dados abertos e interoperáveis. A ausência de diretrizes unificadas, a escassez de investimentos em infraestrutura semântica e a predominância de modelos tradicionais centrados em metadados bibliográficos dificultam a transição para arquiteturas abertas e conectadas. Além disso, a manutenção de sistemas legados baseados em padrões como o *MARC 21 Format for Bibliographic Data*¹⁴ (MARC21) — embora ainda amplamente utilizados — impõe limitações na representação de objetos museológicos ou arquivísticos, reduzindo a riqueza semântica das descrições e comprometendo sua reutilização em contextos ampliados (Wijesundara & Sugimoto, 2018, pp.58-80).

Diante desse contexto, este artigo propõe uma reflexão teórica sobre os fundamentos da interoperabilidade entre acervos do patrimônio cultural, com base na literatura científica da Ciência da Informação, das Humanidades Digitais e das tecnologias da informação aplicadas ao setor *Galleries, Libraries, Archives and Museums*¹⁵ (GLAM). Parte-se do reconhecimento da diversidade documental como valor e desafio, e avança-se na análise crítica de modelos, padrões e estratégias de integração que vêm sendo adotadas em escala internacional. A abordagem adota uma perspectiva exploratório-bibliográfica, orientada à construção de um referencial conceitual que possa subsidiar o desenvolvimento de estudos empíricos e de protótipos voltados à realidade brasileira.

Além de mapear as principais soluções tecnológicas contemporâneas — como destaque para modelos como RDF, *Europeana Data Model*¹⁶ (EDM), Bibliographic Framework Initiative¹⁷ (BIBFRAME), *CIDOC Conceptual Reference Model*¹⁸ (CIDOC CRM) e as ontologias setoriais — o artigo discute a centralidade dos metadados como infraestrutura semântica da curadoria digital, assim como as tensões entre a padronização e a flexibilidade na descrição dos objetos culturais. Também se analisa o papel dos identificadores persistentes¹⁹ (URIs), dos vocabulários controlados e das infraestruturas de Linked Data na criação de redes de significados que ultrapassam as fronteiras institucionais.

Outro aspecto enfatizado é a importância da interdisciplinaridade e da valorização do conhecimento tácito dos profissionais que atuam na mediação, curadoria e preservação dos acervos. A construção de ambientes digitais integrados exige não apenas competências técnicas, mas também sensibilidade para lidar com os contextos históricos, sociais e simbólicos que permeiam os objetos culturais. Assim, modelos centrados na simbiose humano-máquina, como os propostos por O'Neill e Stapleton (2022, pp. 891-903), apontam caminhos para soluções mais responsivas, adaptativas e sustentáveis, colocando o humano no centro das decisões tecnológicas.

Em síntese, este trabalho se insere no debate contemporâneo sobre a curadoria digital e a representação da informação no ecossistema de dados culturais, propondo uma base teórica para pensar a interoperabilidade não apenas como conectividade, mas como um projeto epistêmico e político de construção de redes de memória mais inclusivas, abertas e interligadas.

2. Fundamentação Teórica

2.1 Diversidade Documental e Fragmentação dos Acervos

A diversidade documental constitui um dos maiores valores e, simultaneamente, desafios para as instituições de memória histórica e cultural – bibliotecas, arquivos, museus e centros de documentação. Esses ambientes custodiam conjuntos heterogêneos de objetos, de manuscritos e impressos a artefatos, registros sonoros, imagens, documentos audiovisuais, objetos digitais nativos e patrimônio imaterial. Essa multiplicidade, longe de ser apenas uma característica de acervo, expressa a riqueza da experiência humana e das formas plurais de produção, transmissão e preservação do conhecimento (Marcondes, 2019, pp. 45-52).

Contudo, a pluralidade tipológica e a autonomia histórica dos domínios (biblioteconômico, arquivístico, museológico), decorrem, em grande parte, de um

[...] paradigma custodial, historicista, patrimonialista e tecnicista a formação prática recebida no espaço acadêmico e, sobretudo, no espaço institucional ou funcional das bibliotecas, arquivos e museus apontava claramente para a sobrevalorização da custódia ou guarda, da conservação e do suporte; para a ênfase na ideia de serviço/missão custodial, tendo em vista a salvaguarda da cultura “erudita”; para a hipertrofia da memória como fonte legitimadora da cultura como reforço identitário do mesmo Estado; para a incontornável e crescente importância do acesso ao “conteúdo” (legado tecnicista de Paul Otlet e Henri La Fontaine); e para um arraigado espírito corporativo que fomenta a confusão entre profissão e ciência (SILVA, 2013, pp. 85-103).

Esses fatores resultaram na adoção de modelos descritivos, vocabulários, normas e plataformas técnicas próprias, muitas vezes sem comunicação entre si. O resultado é a formação dos chamados silos de dados: bases e sistemas fechados, criados para finalidades institucionais específicas, que dificultam a interoperabilidade e a descoberta contextualizada de objetos em ambientes digitais (Hyvönen, 2023, pp. 729–744).

Como destaca Hyvönen (2023, pp. 729–744), “conteúdos do patrimônio cultural são fortemente interligados, mas publicados em silos de dados heterogêneos e distribuídos localmente, o que dificulta seu aproveitamento em escala global”. Essa fragmentação se expressa em diversos níveis:

- Modelos de dados incompatíveis, com uso de padrões diferentes - ex: MARC21, *Encoded Archival Description*¹⁹ (EAD), *Dublin Core*²⁰ (DC), CIDOC CRM, FRBRoo²¹;
- Vocabulários e tesouros próprios para cada instituição ou projeto;
- Plataformas digitais isoladas, sem *Application Programming Interface*²² (APIs) ou interoperabilidade com outras bases;
- Barreiras linguísticas e tecnológicas, limitando o compartilhamento transnacional e a reutilização dos acervos;
- Interfaces voltadas ao usuário humano, mas não à disponibilização dos dados para análises computacionais ou integração automatizada.

Do ponto de vista técnico, a padronização e a interoperabilidade de dados culturais foram historicamente secundárias frente às demandas internas de catalogação e gestão. Como observa O’Neill & Stapleton (2022, pp. 891–903), “operar dentro de um silo de dados fechado ou mesmo em uma rede de silos conectados era ‘adequado ao seu tempo’, mas não é o caminho do futuro”. O desafio atual, segundo os autores, é “abrir os dados, adotando padrões comuns e uma mentalidade de mundo aberto, voltada à catalogação para a web semântica”.

A manutenção dos silos institucionais resulta em diversas consequências práticas e epistemológicas:

- Redundância e retrabalho, com desenvolvimento paralelo de vocabulários, esquemas e sistemas semelhantes em diferentes domínios;
- Dificuldade de descoberta e análise transversal: usuários e pesquisadores precisam acessar múltiplos sistemas, aprender diferentes padrões e interfaces, e muitas vezes não conseguem “navegar” entre acervos relacionados (Hyvönen, 2023, pp. 729–744);
- Custo elevado para projetos de integração, já que cada novo consórcio, portal ou iniciativa de mashup demanda esforços de harmonização e mapeamento semânticos entre esquemas;
- Risco para a memória coletiva, pois a incapacidade de interligar fontes distintas impede narrativas multivocais e a reinterpretação crítica do patrimônio cultural em escala global.

A literatura aponta que o problema não é exclusivamente técnico. O modelo tradicional de publicação de dados culturais favorece uma “lógica do local”, na qual cada instituição pública para seu público imediato, sem considerar o potencial de reutilização, enriquecimento e análise global dos dados. Como resultado, há muito trabalho redundante e custoso na criação dos silos, por exemplo, no desenvolvimento de vocabulários, serviços de dados e interfaces (Hyvönen, 2023, pp. 729–744).

Para superar esse cenário, movimentos internacionais como Europeana, *Digital Public Library of America* (DPLA) e a série de portais “Sampo”²³ (na Finlândia) vêm propondo modelos de agregação e publicação baseados em Linked Open Data (LOD) e na utilização de padrões abertos como *Dublin Core*, CIDOC CRM, FRBRoo, *Simple Knowledge Organization System*²⁴ (SKOS), *Web Ontology Language*²⁵ (OWL), RDF, com integração de tesouros internacionais como *Art & Architecture Thesaurus*²⁶ (AAT), *Thesaurus of Geographic Names*²⁷ (TGN), *Union List of Artist Names*²⁸ (ULAN), APIs e pontos de acesso SPARQL²⁹. Essas iniciativas buscam promover não só a agregação de dados, mas também o enriquecimento mútuo dos conteúdos, a publicação orientada a dados e o desenvolvimento de ambientes colaborativos para produção, curadoria e análise do patrimônio cultural em escala global (Hyvönen, 2023, pp. 729–744).

A fragmentação dos acervos, portanto, não é apenas um problema operacional, mas epistemológico, social e político, impactando a construção de narrativas sobre a memória coletiva, o direito ao acesso à informação e o potencial de uso educativo e científico dos acervos digitais.

2.2 Metadados, Modelos de Integração e Infraestruturas Semânticas

A integração de acervos culturais digitais representa um dos maiores desafios contemporâneos para bibliotecas, arquivos e museus, especialmente em razão da diversidade de padrões de metadados, modelos descritivos e infraestruturas técnicas adotadas historicamente. O advento das tecnologias da Web Semântica e dos Dados Abertos Interligados (*Linked Open Data* – LOD) inaugurou um novo paradigma, no qual o papel dos metadados - definidos como infraestruturas descritivas da representação da informação - , transcende a mera descrição estática e passa a ser fundamental para a descoberta, interconexão e reutilização dos objetos culturais em escala global.

Como apontam Silva e Terra (2024, pp. 93-107, tradução nossa), “a Web Semântica permite que os dados sejam vinculados na Web e estrutura a informação para uso por humanos e máquinas. Além disso, torna explícitas as relações entre os dados, possibilitando a criação de Linked Data”. No contexto das instituições GLAM, a transição de modelos tradicionais para padrões abertos, ontologias e infraestruturas compatíveis com a Web Semântica tornou-se uma condição necessária para romper com os “silos de dados” e promover ambientes verdadeiramente colaborativos e inovadores.

Gaitanou et al. (2024, pp. 204-244, tradução nossa) reforçam esse argumento ao afirmar que

Há um forte contraste entre a Web e o principal catálogo das instituições culturais que existem como silos e isolados do ecossistema mais amplo da Web. Apesar dos esforços, a interoperabilidade de dados ainda é uma questão em aberto, que se relaciona com a necessidade de converter os dados existentes em dados compatíveis com a Web Semântica.

O desenvolvimento de modelos conceituais mais abertos e de infraestruturas semânticas compartilhadas tem permitido avanços importantes em projetos internacionais, como Europeana, DPLA e Sampo, ainda que persistam desafios de harmonização, qualidade dos metadados, governança e formação de equipes multidisciplinares. Neste contexto, a compreensão das transformações dos padrões de metadados, das arquiteturas LOD e dos modelos de integração semântica torna-se fundamental para a construção de ecossistemas informacionais abertos, acessíveis e socialmente relevantes.

2.2.1 Diversidade de Padrões e Modelos

A integração de acervos digitais de arquivos, bibliotecas e museus apresenta um desafio intrínseco relacionado à diversidade de padrões, modelos descritivos e vocabulários adotados historicamente por cada domínio. Como destaca Marcondes (2016, pp. 61-83),

Diferenças de gestão e curadoria de seus acervos, diferenças na conceituação dos objetos dos acervos a descrever (diferentes ‘objetos’ a representar: livros, fundos e séries, objetos museológicos únicos), idiosincrasias relacionadas a tradições profissionais distintas, padrões diferentes, etc., tornam a integração dos acervos dessas instituições um desafio conceitual e tecnológico.

Esse panorama de pluralidade se manifesta, principalmente, nas escolhas de modelos de metadados e normas técnicas. Segundo Sacramento et al. (2015, pp. 1-49), “Os acervos de natureza física de tais áreas – como livros de prateleira, documentos impressos, obras de arte, objetos de museu e, até mesmo, micro-organismos em lâminas de microscópio, podem ter representantes digitais que permitam a perpetuação dos objetos desses acervos [...]” e, para viabilizar a gestão integrada,

O uso de metadados específicos de domínio, de preservação e de interoperabilidade apresentam-se como um dos mecanismos essenciais para apoiar a construção ou a adoção de ferramentas tecnológicas que viabilizem a preservação e a integração de objetos digitais pertencentes aos acervos em questão. Além disso, a natureza da Web como um espaço sem limitações geográficas, permite que esses representantes (ou objetos) digitais sejam consultados em qualquer parte do mundo (Sacramento et al., 2015, pp. 1-49).

Entre os padrões amplamente utilizados, destacam-se o MARC21 para registros bibliográficos, EAD e *General International Standard Archival Description*³⁰ - ISAD(G) para descrição arquivística e o VRA Core³¹ para museus, cada qual com estrutura, granularidade e escopo próprios. Conforme O’Neill e Stapleton (2022, pp. 891-903),

O processo de digitalização do patrimônio cultural requer a quebra de silos em várias áreas – uma delas é a das disciplinas acadêmicas, para possibilitar trabalhos interdisciplinares ricos. Isso estabelece as bases para a emancipação de outro tipo de silo, os silos de conhecimento, tanto tradicional quanto digital, mantidos em instituições como galerias, bibliotecas, arquivos e museus. Quebrar os silos disciplinares é a chave para abrir os silos institucionais de conhecimento.

Como já discutido por Silva (2013, pp. 85-103), a predominância de um paradigma custodial, patrimonialista e tecnicista, historicamente forjou diferentes perspectivas sobre o “objeto de memória” e a missão das instituições de patrimônio, influenciando diretamente a adoção e fragmentação dos padrões descritivos nos domínios GLAM. Além disso, a coexistência desses

diferentes padrões reflete não apenas escolhas técnicas, mas também distintas tradições epistemológicas e históricas sobre o que constitui o “objeto de memória” em cada domínio do patrimônio. Buckland (2006, pp. 3-14) enfatiza que “toda a descrição é uma atividade de linguagem [...], sempre e necessariamente de base cultural, pois descrições são baseadas em conceitos, definições e compreensões que têm desenvolvido em uma comunidade”. Além disso, o próprio “conceito de metadados pode variar de comunidade para comunidade. [...] A Ciência da Informação classifica os metadados em descritivos, estruturais e administrativos, mesmo não havendo um consenso na área” (Sacramento et al., 2015, pp. 1-49). Assim, a fragmentação dos modelos de descrição e representação resulta tanto da pluralidade institucional quanto das escolhas históricas e culturais dos campos do patrimônio.

Contudo, há padrões emergentes, como o Dublin Core, *Europeana Data Model* (EDM) e BIBFRAME, que buscam viabilizar a interoperabilidade por meio de descritores genéricos e flexíveis. Ribeiro et al. (2020, pp. 247-264) destacam que “o uso das tecnologias da Web Semântica e dos princípios *Linked Data* tem sido apontado como facilitador para interligar dados e recursos oriundos de diferentes acervos e procedências a partir do momento em que possibilita que se estabeleçam metadados de ligações semânticas em contextos específicos”.

Porém, o desafio persiste: como harmonizar modelos que, além de distintos, muitas vezes apresentam sobreposições, lacunas ou incompatibilidades? Buckland (2006, pp. 3-14) enfatiza que

Os metadados possuem dois componentes: um formato e um conjunto de valores. XML^[32], Dublin Core e os registros de catálogo de biblioteca MARC são formatos conhecidos e são associados a padrões específicos para especificar os tipos de descrições que podem ser utilizados com eles. [...] Entretanto, é sempre fortemente recomendado que metadados descritivos sigam formas padronizadas, por exemplo, utilizar um formato padrão e terminologia amplamente utilizada. O uso de formatos padronizados para armazenamento e exibição torna o uso de metadados mais fácil. O uso de vocabulários padrão possui a vantagem da consistência e auxilia na compreensão.

A discussão recente também passa pela necessidade de modelos de integração que operem “acima” dos padrões específicos, como as ontologias de domínio, os crosswalks e as arquiteturas baseadas em *Linked (Open) Data* (LOD) e Web Semântica. Como ressalta Zeng (2019) apud Marcondes (2023, pp. 85-95) “coleções digitais de arquivos, bibliotecas e museus são insumos importantes para a pesquisa em humanidades digitais [...]. Tais coleções, quando publicadas como LOD, são abertas, não constituem silos de informação. Seus objetos digitais podem ser utilizados por qualquer pessoa e interligados com outros recursos”.

Portanto, o avanço em direção à integração efetiva dos acervos culturais depende não só de soluções técnicas, mas de abordagens colaborativas e de uma visão crítica sobre a história dos padrões, a cultura institucional e a abertura para revisitar paradigmas tradicionais de organização da informação.

2.2.2 Das fichas descritivas ao Linked Data

Modelos e esquemas detalhados para descrever recursos de informação existem há muito tempo, mas não são compatíveis com os padrões da Web, muito menos com os princípios da Web Semântica. Historicamente, a integração de acervos digitais foi fundada na troca de metadados entre instituições, fundamentada em protocolos e formatos como MARC21, *Dublin Core*, *Metadata Object Description Schema*³³ (MODS), *Metadata Encoding & Transmission Standard*³⁴ (METS), EAD e Open Archives Initiative Protocol for Metadata Harvesting³⁵ (OAI-PMH). Nessa abordagem, os metadados funcionavam como “fichas descritivas”, criadas para cada objeto e exportadas em lotes para sistemas agregadores ou bases centrais (Silva e Terra, 2024, pp. 93-107). Esse modelo

tradicional, centrado na descrição e no intercâmbio de registros, consolidou avanços importantes, mas permaneceu limitado por barreiras técnicas e semânticas (Marcondes (2016, pp. 61-93).

As limitações dessa abordagem são múltiplas. Metadados são mais do que simples descrições. Eles são parte da infraestrutura necessária para garantir que objetos digitais sejam encontrados, compreendidos e reutilizados (Buckland, 2006, pp. 3-14). Entretanto, os formatos tradicionais operam, em geral, sem explicitação de relações complexas entre entidades, o que restringe a possibilidade de descoberta transversal e o reuso inteligente dos dados.

O paradigma da troca de metadados representou um grande avanço para a interoperabilidade, mas tornou-se insuficiente diante da demanda por integração dinâmica, enriquecimento automático, descoberta transversal e contextualização dos acervos. O paradigma do Linked Data surge, assim, como resposta a essas limitações, promovendo relações explícitas, uso de ontologias e publicação aberta e interligada de dados.

Neste sentido, a emergência do paradigma dos Dados Abertos Interligados (*Linked Open Data* – LOD) representa uma inflexão significativa. Não se trata apenas de uma evolução técnica, mas de uma transformação cultural e epistêmica no modo como os metadados são produzidos, publicados e conectados. Como destaca Silva e Terra (2024, pp. 93-107), “a necessidade de evoluir os formatos tradicionais de metadados utilizados em bibliotecas, museus e arquivos para modelos *Linked Data* desencadeou iniciativas em várias instituições, levando ao surgimento de diferentes modelos conceituais para cada uma dessas áreas”. Nesse novo cenário, o metadado deixa de ser uma “ficha” para se tornar um nó em um grafo global, interligado por *URIs*³⁴ persistentes e relações formalizadas em ontologias abertas e compartilhadas.

O salto qualitativo está justamente na capacidade do *Linked Data* de criar redes semânticas, tornando explícitas as relações entre diferentes entidades (obras, autores, lugares, eventos), conectando os acervos a contextos mais amplos e promovendo novas formas de descoberta e análise (Ribeiro et al., 2020, pp. 247-264). Diversas bibliotecas e organizações de padronização desenvolveram modelos conceituais e padrões de metadados adequados para a *Web* e, particularmente, para ambientes de Dados Conectados. O *Europeana Data Model*, por exemplo, foi desenvolvido para acomodar as necessidades de um ambiente de dados abertos, em que recursos podem ser descritos por múltiplos parceiros e suas relações explicitadas e conectadas a recursos externos, como ontologias e bases de autoridade (Silva & Terra, 2024, pp. 93-107).

Esse novo ecossistema potencializa não só o reuso, mas também o enriquecimento e a validação dos dados por múltiplas comunidades e aplicações. Como afirmam Salse et al. (2024, pp. 447-495, tradução nossa),

o desenvolvimento de esquemas de troca de metadados [como Dublin Core, Darwin Core ou LIDO^[36]], a criação de repositórios digitais que coletam conteúdos de diferentes fontes (como Europeana, DPLA ou GBIF) e a presença de protocolos de agregação e transferência (como OAI-PMH) parecem sustentar a ideia de que a cooperação é hoje o caminho preferencial.

No entanto, o verdadeiro salto do *Linked Data* está na superação das limitações impostas pelos antigos silos institucionais. A literatura de Marcondes (2016, pp. 61-93; 2019, pp. 45-52; 2022, pp. 1-13; 2023, pp. 85-95) demonstra que a *Web Semântica* — ao utilizar ontologias, grafos de conhecimento e padrões abertos como *RDF* e *Linked Open Data* — proporciona meios avançados de representação, conexão e organização da informação, permitindo que dados de diferentes acervos sejam publicados de maneira interligada, interoperável e reutilizável. Essa abordagem supera os limites da agregação tradicional de registros ao viabilizar a articulação semântica entre múltiplas fontes, a automação na descoberta e o enriquecimento dinâmico dos acervos digitais.

Isso impacta não apenas a governança, com novos arranjos para colaboração e gestão distribuída, mas também a própria lógica de produção e circulação do conhecimento, promovendo a ciência aberta, a transparência e a democratização do acesso ao patrimônio cultural.

Quadro 1: Comparativo – Troca de Metadados vs. Linked Data

Dimensão	Paradigma da Troca de Metadados	Paradigma do Linked Data
Objetivo	Compartilhar registros descritivos para integração básica entre sistemas ou agregadores	Interligar recursos, entidades e contextos por meio de redes semânticas abertas e colaborativas
Modelo de Dados	Estrutura tabular, baseada em registros isolados, campos fixos e hierárquicos; formatos estruturados (planilha, XML, banco de dados relacional)	Grafo de conhecimento; triplas RDF; uso de URIs e ontologias abertas e flexíveis
Padrões principais	MARC21, EAD, MODS, METS Dublin Core, (padrões de descrição e exportação estruturada)	RDF, OWL, SKOS, EDM, CIDOC CRM, BIBFRAME, LIDO (padrões de modelagem semântica e interoperável)
Interoperabilidade	Predominantemente sintática/estrutural; baseada em crosswalks e mapeamentos pontuais entre formatos	Semântica explícita; interoperabilidade nativa entre domínios, sistemas e plataformas
Relações entre Objetos	Relações implícitas, codificadas em campos textuais ou pouco formalizadas	Relações explícitas e formalizadas entre entidades, eventos, lugares, agentes
Enriquecimento de Dados	Limitado, geralmente manual ou centralizado, dependente de atualização de registros e novos envios	Enriquecimento dinâmico e automático, colaborativo, via links externos e autoridades globais
Descoberta e Navegação	Busca restrita ao próprio repositório ou agregador; baixa capacidade de inferência	Descoberta global, navegação transversal, inferência semântica e integração em escala

Atualização e Curadoria	Atualização cíclica, via novos envios de registros ou ingestões periódicas	Curadoria distribuída, atualização contínua, colaboração entre instituições e usuários
Papel dos Profissionais	Foco em catalogação, normalização e controle de formatos; ênfase em manutenção de registros e padrões internos	Protagonismo na modelagem conceitual, curadoria semântica, mediação entre comunidades e ecossistemas
Exemplos de Implementação	Bibliotecas digitais via OAI-PMH; Catálogos MARC21; Repositórios baseados em Dublin Core, MODS ou EAD; portais nacionais e locais de metadados	Europeana, DPLA, Wikidata ³⁷ , portais Sampo, repositórios LOD, catálogos BIBFRAME, sistemas baseados em EDM

Fonte: Elaborado pelo autor (2025)

O quadro acima sintetiza as principais diferenças entre o paradigma tradicional de troca de metadados e o paradigma do *Linked Data* na integração de acervos digitais. A comparação evidencia como a evolução dos modelos, padrões e práticas impacta diretamente a interoperabilidade, o enriquecimento dos dados e o papel dos profissionais da informação, destacando a transição de abordagens isoladas e centralizadas para ecossistemas colaborativos, dinâmicos e semanticamente interligados.

Neste sentido, a transição das fichas descritivas para o paradigma do Linked Data representa um avanço qualitativo fundamental: desloca-se de descrições isoladas e centralizadas para a construção de redes semânticas dinâmicas, abertas e colaborativas, que favorecem o reuso, o enriquecimento e a inovação nas práticas contemporâneas de curadoria digital.

2.2.3 Modelos de Integração: Exemplos e Desafios

A busca pela integração efetiva de acervos culturais digitais tem levado ao desenvolvimento de diferentes modelos e plataformas em escala internacional. Tais modelos são essenciais para viabilizar a interoperabilidade entre bases heterogêneas, promover o acesso transversal e fomentar a reutilização e o enriquecimento dos acervos. Assim, a consolidação de ecossistemas digitais abertos no patrimônio cultural internacional resultou do desenvolvimento de modelos de integração que vão além da mera agregação de registros, propondo arquiteturas baseadas em *Linked Data*, grafos de conhecimento e interoperabilidade semântica. Destacam-se, nesse cenário, iniciativas como a *Europeana*, a *Digital Public Library of America* (DPLA) e os *portais Sampo*, que oferecem referências paradigmáticas para o campo.

O *Europeana Data Model* (EDM) tornou-se um marco na integração de acervos ao estabelecer uma arquitetura semântica flexível e orientada à Web Semântica. O EDM não é construído sobre qualquer padrão comunitário específico, mas adota uma estrutura aberta, baseada na Web Semântica, capaz de acomodar a variedade e a riqueza dos padrões de diferentes comunidades (Isaac, 2013, pp. 1-35). O EDM possibilita a representação de múltiplos níveis de descrição, preservando a diversidade dos dados originais enquanto os torna semanticamente interoperáveis.

A capacidade do EDM de criar uma “rede de objetos de patrimônio cultural digitais e nativos digitais, graças à representação de links e relações semânticas entre objetos” (Charles & Isaac, 2015, pp. 1-21, tradução nossa), exemplifica o salto qualitativo em relação aos modelos tradicionais. O EDM apoia tanto a descrição detalhada de cada objeto quanto o enriquecimento automático com dados externos, integrando ontologias, tesouros, listas de autoridade e fontes como SKOS, DBpedia³⁸ e GeoNames³⁹.

No contexto americano, a *Digital Public Library of America* (DPLA) adota um perfil de aplicação próprio — o *DPLA Metadata Application Profile* (MAP⁴⁰) — que,

[...] permite a integração de metadados criados ou agregados e compartilhados por nossos hubs em vários padrões de metadados. Isso nos permite representar e compartilhar adequadamente esses dados e enriquecê-los para maior descoberta e acesso. O DPLA MAP é um perfil de aplicativo ou um conjunto de elementos de metadados, retirados de vários esquemas para um uso local específico. É também um modelo de metadados semânticos, ou uma estrutura abstrata que descreve as relações entre diferentes tipos de dados sobre a mesma coisa” (DPLA, 2017, tradução nossa).

O projeto demonstra como o alinhamento de centenas de parceiros institucionais, cada qual com sistemas e padrões diferentes, pode ser viabilizado pela definição de mapeamentos semânticos precisos e pela adoção de um grafo comum.

Outro exemplo relevante é o *Sampo Model*, implementado em portais finlandeses como o *WarSampo*, que utiliza tecnologias da Web Semântica para possibilitar a publicação, o enriquecimento e a análise de dados interligados do patrimônio cultural utilizando tecnologias da web semântica e design de portal centrado no usuário. O *Sampo Model* estabelece princípios fundamentais para a publicação, o enriquecimento e a análise de dados interligados do patrimônio cultural utilizando tecnologias da web semântica e design de portal centrado no usuário. Esses princípios incluem: apoio à criação colaborativa e publicação de dados; uso de uma infraestrutura ontológica aberta e compartilhada; distinção clara entre o serviço LOD e a interface do usuário; oferta de múltiplas perspectivas para os mesmos dados; padronização do uso do portal por meio de um ciclo simples de filtragem-análise em dois passos; e apoio à análise de dados e descoberta de conhecimento, além da mera exploração dos dados (Hyvönen, 2023, pp. 729-744).

Ainda No contexto dos modelos de integração orientados ao ecossistema LOD, o BIBFRAME tem ganhado destaque ao proporcionar bases semânticas robustas para a representação e compartilhamento de dados bibliográficos em escala global. Mantido pela US Library of Congress, o BIBFRAME foi concebido como evolução do MARC, visando “[...] estabelecer as bases para o futuro da descrição bibliográfica, sendo uma evolução das iniciativas da própria Biblioteca do Congresso desde o Projeto MARC nos anos 1970”, (Marcondes, 2023, pp. 85-95, tradução nossa).

Um de seus diferenciais é a modelagem nativa em RDF, tornando “[...] trivial o mapeamento da entidade bibliográfica descrita em um registro MARC — o Sujeito — de uma tripla RDF, as propriedades identificadas pelos parágrafos/códigos de subcampo do MARC — o Predicado da tripla —, e o conteúdo de tais parágrafos/códigos de subcampo — seu Objeto.” (Marcondes, 2023, p. 88). A estrutura do BIBFRAME organiza-se em três níveis de abstração — Work, Instance e Item — “[...] simplificando de certa forma os quatro níveis de abstração dos modelos Functional Requirements for Bibliographic Records⁴¹ (FRBR) e CRM”, (Marcondes, 2023, p. 88, tradução nossa). A maioria das propriedades fornecidas pelo BIBFRAME diz respeito a “[...] propriedades descritivas de entidades bibliográficas ou relacionamentos entre entidades de diferentes níveis de abstração” (Marcondes, 2023, pp. 85-95).

O BIBFRAME, assim como o EDM e outros modelos contemporâneos, contribui para a superação das limitações dos padrões legados, favorecendo a integração, interoperabilidade e enriquecimento semântico dos acervos bibliográficos em escala internacional.

Já o IFLA⁴² *Library Reference Model*⁴³ (IFLA LRM) representa a evolução e a harmonização dos modelos conceituais da família FR (FRBR, FRAD e FRSAD), consolidando uma estrutura única para descrição de dados bibliográficos no contexto das bibliotecas. O IFLA LRM é “um modelo conceitual de alto nível e, como tal, pretende ser um guia ou base para formular regras de catalogação e implementar sistemas bibliográficos” (Riva, Bœuf & Žumer, 2017a, p. 09 apud Arakaki, 2020, pp. 163-186). Sua proposta é fornecer uma estrutura de referência capaz de analisar metadados não administrativos, sendo fiel às entidades, aos relacionamentos e às limitações de cardinalidade estabelecidos no modelo .

A principal inovação do IFLA LRM está na harmonização dos modelos conceituais FRBR, *Functional Requirements for Authority Data*⁴⁴ (FRAD) e *Functional Requirements for Subject Authority Data*⁴⁵ (FRSAD), resultando no modelo IFLA LRM, o qual busca simplificar entidades e tarefas do usuário e fornecer uma base única para modelagem conceitual de registros bibliográficos (Arakaki, 2020). Estruturalmente, o modelo adota como superclasse a entidade ‘Res’, englobando ‘Obra’, ‘Expressão’, ‘Manifestação’, ‘Item’, ‘Agente’, ‘Nomen’, ‘Lugar’ e ‘Intervalo de Tempo’, definindo uma hierarquia clara de classes e subclasses conectadas por relacionamentos (Arakaki, 2020, pp. 163-186). O IFLA LRM se destaca, ainda, por alinhar-se aos princípios do Linked Data e tecnologias semânticas (RDF, OWL, SPARQL), possibilitando a instanciação das entidades e atributos em ambientes interoperáveis, abrindo caminho para catálogos digitais mais abertos, conectados e eficientes.

O CIDOC *Conceptual Reference Model* (CIDOC CRM), desenvolvido pelo Comité *International pour la Documentation* do Conselho Internacional de Museus⁴⁶ (ICOM-CIDOC), tornou-se padrão ISO 21127 para modelagem conceitual no domínio do patrimônio cultural. Segundo a norma seu principal objetivo,

[...] é oferecer uma base conceitual para a integração, mediação e intercâmbio de informações entre organizações de patrimônio cultural, como museus, bibliotecas e arquivos. Este documento visa fornecer um ponto de referência comum contra o qual fontes divergentes e incompatíveis de informação possam ser comparadas e, em última instância, harmonizadas. (ISO 21127:2023, p. v, tradução nossa)

Diferente de modelos restritos à descrição de objetos, o CIDOC CRM constitui uma ontologia que fornece uma representação formal do esquema conceitual subjacente aos sistemas de documentação e bancos de dados utilizados por instituições de patrimônio cultural. Seu principal papel é viabilizar a troca e integração de informações entre fontes heterogêneas, ao oferecer definições semânticas que tornam possível transformar fontes de informação díspares e localizadas em um recurso global coerente, tanto em contextos institucionais quanto em ambientes digitais abertos (ISO 21127:2023). Segundo Doerr (2003, p. 75-92, tradução nossa),

O CIDOC CRM é uma ontologia de alto nível para possibilitar a integração de informações de dados do patrimônio cultural e sua correlação com informações de bibliotecas e arquivos. [...] O CIDOC CRM analisa as conceitualizações comuns por trás das estruturas de dados e metadados para apoiar a transformação, mediação e fusão de dados.

Diferentemente de modelos voltados apenas à descrição de entidades ou itens isolados, o CIDOC CRM é projetado para permitir a descrição de relações complexas entre objetos, eventos, atores e contextos históricos, favorecendo uma contextualização rica e significativa dos acervos. Segundo

Marcondes (2021), o modelo é orientado por eventos e busca fornecer informações contextuais essenciais para a compreensão dos objetos de patrimônio cultural, relacionando-os a atores e acontecimentos e permitindo, assim, um entendimento mais aprofundado das conexões históricas e culturais envolvidas. Essa abordagem facilita a interoperabilidade semântica entre bases de museus, arquivos e bibliotecas, viabilizando integrações ricas, contextualizadas e abertas para pesquisa, ensino e difusão do patrimônio cultural.

Apesar desses avanços, permanecem desafios substanciais. O mapeamento de dados existentes para modelos de integração de bases de dados heterogêneas do patrimônio cultural requer análise cuidadosa dos metadados de origem, com a manutenção de contextos e proveniência e pode demandar limpeza e transformação dos dados (harmonização de metadados e vocabulários). Como destacam Lemos, Coelho Junior & Martins, (2023, pp. 1-26) “questões relativas à qualidade de dados frequentemente não são levantadas, considerando os diversos tipos de bancos de dados e sistemas de informação ora envolvidos em processos de organização, modelagem e representação da informação”.

Além disso, a integração requer um alto nível de especialização e capacitação técnica, além de compromisso sobretudo com a superação de barreiras políticas, com a quebra dos silos institucionais. Ou seja, os modelos internacionais de integração demonstram que a construção de ambientes digitais verdadeiramente interoperáveis exige não apenas tecnologias avançadas, mas também a mobilização de comunidades, governança colaborativa e investimento constante em curadoria, controle de qualidade e atualização dos dados.

2.2.4 Infraestruturas Semânticas e Arquiteturas LOD

A efetiva integração e reuso de dados culturais digitais dependem de infraestruturas semânticas robustas, capazes de sustentar a publicação, interligação e enriquecimento contínuo dos acervos em escala global. Infraestruturas semânticas combinam ontologias, vocabulários controlados, grafos de conhecimento, repositórios LOD, APIs e serviços de reconciliação (Também conhecido como vinculação de registros ou correspondência de dados), compondo a espinha dorsal dos ecossistemas digitais abertos.

Conforme sintetiza Silva & Terra (2024, pp. 93-107), “Linked Data baseia-se no modelo de metadados semânticos RDF (Resource Description Framework)”, o que possibilita a publicação e interligação de dados por meio de triplas sujeito-predicado-objeto e identificadores persistentes. O *Europeana Data Model* (EDM), por exemplo, fundamenta sua infraestrutura semântica na utilização de URIs, relações explícitas e no alinhamento a vocabulários internacionais como SKOS, GeoNames, DBpedia, Virtual International Authority File⁴⁷ (VIAF) e outros. Com o EDM, a *Europeana* consegue organizar e articular objetos do patrimônio cultural provenientes de múltiplas instituições, graças ao uso de links semânticos e à explicitação das relações entre esses objetos (Charles & Isaac, 2015, pp. 1-21).

Assim, o uso de ontologias abertas – como o CIDOC CRM no domínio museológico ou a própria estrutura EDM para agregação – é fundamental para viabilizar interoperabilidade em nível semântico. Isso garante que instituições com diferentes tradições e esquemas de metadados possam interoperar e compartilhar seus acervos de forma integrada, sem perder a expressividade ou a riqueza dos dados originais (Isaac et al., 2013, pp. 1-35).

Além disso, a arquitetura LOD pressupõe a utilização de boas práticas de publicação aberta, incluindo a atribuição de URIs descritivas, a disponibilização dos dados em formatos legíveis por máquina (como RDF/XML, *Turtle*⁴⁸, JSON-LD⁴⁹) e o estabelecimento de links entre conjuntos de dados distintos. Essa arquitetura permite que diferentes repositórios e coleções compartilhem e

conectem seus dados de maneira dinâmica, promovendo o enriquecimento semântico e facilitando a descoberta transversal de recursos culturais.

O sucesso dessas infraestruturas depende, ainda, da adoção de serviços de reconciliação e enriquecimento, que permitem conectar entidades locais a referências globais (autoridades, lugares, conceitos), reduzindo ambiguidades e ampliando o potencial de descoberta. O EDM, por exemplo, possibilita o enriquecimento de dados a partir de fontes autorizadas, incluindo arquivos de autoridade e tesouros de outros provedores ou editores. Esse processo permite, o alinhamento a vocabulários de autoridade (pessoas, tipos de objetos, lugares, períodos) e recursos como a DBpedia ou Wikidata, além de contribuir para identificar registros duplicados e traduzir metadados por associação (Europeana Foundation, n.d.).

Por fim, a consolidação de infraestruturas semânticas e arquiteturas LOD implica desafios recorrentes de governança, atualização, sustentabilidade e qualidade dos dados. O processo de mapeamento e harmonização de metadados é complexo e requer alto nível de especialização, além de compromisso institucional com curadoria contínua e controle de qualidade. A construção de infraestruturas semânticas e arquiteturas LOD é condição indispensável para uma curadoria digital inovadora, inclusiva e aberta, capaz de integrar múltiplos acervos, enriquecer contextos e promover novas formas de apropriação e reuso do patrimônio cultural no ambiente digital contemporâneo.

2.2.5 Tendências e Desafios

O avanço dos ecossistemas de integração de acervos culturais digitais aponta para um cenário em que princípios, tecnologias e práticas inovadoras tornam-se fundamentais para garantir a efetividade, sustentabilidade e impacto social dessas infraestruturas. Entre as principais tendências, destaca-se a adoção dos Princípios FAIR — *Findable, Accessible, Interoperable, Reusable* — que representam um conjunto conciso e mensurável de diretrizes projetadas para melhorar fundamentalmente a infraestrutura de dados, facilitando a descoberta e o reuso de dados acadêmicos, científicos e objetos de pesquisa digital. De fato, os princípios FAIR são considerados um modelo essencial para a gestão contemporânea de dados científicos e culturais, uma vez que garantem que os dados sejam efetivamente encontráveis, acessíveis, interoperáveis e reutilizáveis, potencializando a integração e o compartilhamento de informações entre instituições culturais digitais (Wilkinson et al., 2016). Isso pode ser estendido à garantia de que os próprios dados digitais sejam considerados e preservados como parte do próprio patrimônio, com requisitos adicionais para a relevância, confiabilidade e integridade (Hermon & Niccolucci, 2021, pp. 251-255).

Outra tendência crucial é a automação dos processos de enriquecimento e reconciliação de metadados, com a incorporação de técnicas de inteligência artificial (IA) para desambiguação, linkagem semântica e detecção de relações ocultas entre entidades. Avanços recentes em tecnologia da informação e inteligência artificial oferecem oportunidades notáveis para o enriquecimento automático dos metadados do patrimônio cultural com recursos mínimos. Embora estes avanços em tecnologia da informação e inteligência artificial ofereçam oportunidades notáveis para o enriquecimento automático de metadados do patrimônio cultural, a participação humana nesse processo continua sendo fundamental. Como observa Kaldeli (2023), a estratégia mais eficaz envolve justamente a combinação de algoritmos com validação humana, promovendo uma abordagem híbrida que potencializa tanto a eficiência da automação quanto a precisão e o contexto aportados pelos especialistas. Ou seja, "[...] combina algoritmos com validação humana para o enriquecimento de metadados de patrimônio cultural em larga escala" (Kaldeli, 2023, tradução nossa).

A governança e a qualidade dos dados emergem como preocupações centrais, exigindo modelos colaborativos, políticas de atualização permanente e mecanismos de validação distribuída. Soma-se a isso a fragmentação institucional e a ausência de diretrizes unificadas, pois a sustentabilidade e a eficácia das infraestruturas LOD dependem de políticas compartilhadas, curadoria contínua e participação ativa das comunidades envolvidas. Bettencourt & Marcondes (2019, pp. 44-61) destacam que a consolidação de políticas sustentáveis para integração de acervos digitais, sobretudo no Brasil, ainda enfrenta a “proverbial dificuldade das instituições brasileiras de articularem-se em projetos interinstitucionais, para além de suas atribuições rotineiras”. Para superar tal fragmentação, os autores defendem a adoção de políticas nacionais com sustentação política e econômica permanente, fundamentadas na cooperação, governança compartilhada e participação ativa das comunidades envolvidas.

Manter a qualidade e a confiança dos dados em ambientes abertos e distribuídos permanece um desafio contínuo, que requer novas formas de cooperação institucional e soluções técnicas.

Embora nos últimos anos instituições do patrimônio histórico e cultural aderirem a processos de digitalização e disponibilização de conteúdos na Web, proporcionando acesso a conteúdos e democratizando o conhecimento científico e cultural à sociedade, como destacam Lemos, Coelho Júnior & Martins (2023, pp.1-26), apesar desses esforços,

[...] investir somente na digitalização de objetos culturais não é suficiente, visto que questões relativas à qualidade de dados frequentemente não são levantadas, considerando os diversos tipos de bancos de dados e sistemas de informação ora envolvidos em processos de organização, modelagem e representação da informação.

Ainda é escassa a implementação de processos robustos de garantia de qualidade em repositórios digitais, e a ausência de procedimentos transferíveis para outros contextos dificulta a consolidação de ecossistemas interoperáveis e confiáveis, demandando esforços contínuos de avaliação, curadoria e colaboração entre instituições.

Por fim, destaca-se o crescente papel da formação e capacitação profissional. O campo demanda cada vez mais especialistas capazes de atuar na interseção entre curadoria digital, tecnologias semânticas e gestão colaborativa de acervos, promovendo uma atuação crítica, ética e inovadora. O desenvolvimento de competências transversais — desde a compreensão de ontologias e vocabulários, passando pela utilização de ferramentas de IA e reconciliação, até a liderança de processos colaborativos — é condição indispensável para que os profissionais da informação possam liderar a construção de ambientes digitais abertos, interoperáveis, integrados, confiáveis e socialmente relevantes. Assim, o fortalecimento do capital humano deve ser reconhecido como eixo estratégico das políticas de integração e curadoria digital, respondendo à crescente demanda por competências transversais, visão crítica e compromisso social.

Sobre a discrepância entre os avanços internacionais e a realidade brasileira, Siqueira & Martins (2022, pp. 204-224) destacam que,

O cenário internacional nos apresenta diversas soluções para uma interface integrada, sendo a Europeia o exemplo mais conhecido. No Brasil, ainda não existe uma solução amplamente aplicada, pois as iniciativas nacionais geralmente se restringem a pequenos projetos em instituições locais.

Neste contexto, os desafios vão além da mera adoção de tecnologias: envolvem a construção de modelos participativos, iniciativas de políticas públicas e a revisão de políticas institucionais, o fortalecimento da cultura de dados abertos e o investimento contínuo no capital humano que sustenta as práticas de integração e curadoria digital.

3. Metodologia

Este estudo adota uma abordagem qualitativa de natureza exploratória, fundamentada na análise de conteúdo aplicada a uma revisão narrativa da literatura. A opção metodológica visa identificar, organizar e interpretar categorias teóricas relevantes sobre a interoperabilidade entre acervos culturais, com foco na integração de metadados em ambientes digitais heterogêneos. A abordagem é adequada para compreender os fundamentos conceituais da curadoria digital, dos padrões descritivos e das infraestruturas tecnológicas voltadas à integração de dados culturais.

O corpus documental foi selecionado com base em critérios de relevância teórica, atualidade, abrangência temática e reconhecimento na comunidade científica. Foram priorizadas publicações dos últimos 25 anos (2000–2025) aproximadamente, incluindo artigos acadêmicos, relatórios técnicos, documentos institucionais e estudos de caso nacionais e internacionais. As buscas foram realizadas em bases como Scopus, Web of Science, Google Scholar, Portal de Periódicos da CAPES, além de repositórios institucionais e especializados.

A análise de conteúdo foi conduzida conforme Bardin (2011), em três etapas: (1) leitura exploratória e pré-análise, para familiarização com o material; (2) codificação temática, com destaque para as ocorrências de conceitos como silos de dados, metadados, interoperabilidade, integração de bases de dados, modelos de integração, FAIR principles, Linked Data, EDM, CIDOC CRM, RDF, entre outros; e (3) categorização dos achados em eixos analíticos.

As principais categorias identificadas foram: (a) modelos e padrões de metadados; (b) infraestruturas semânticas e web de dados; (c) integração de acervos e estratégias de crosswalks; (d) limitações da governança e qualidade dos dados; e (e) desafios no contexto brasileiro. A triangulação entre os documentos técnicos e os estudos acadêmicos permitiu uma compreensão ampliada das tensões entre padronização, representatividade e aplicabilidade nos sistemas de memória.

Ainda que este não seja um estudo empírico, o procedimento adotado garante rigor analítico na construção de um panorama teórico sólido e crítico, útil para subsidiar futuras iniciativas práticas no campo da curadoria digital e da interoperabilidade semântica. A explicitação dos critérios de seleção, a diversidade das fontes e a estruturação categorial contribuem para a validade dos achados.

4. Resultados Parciais e Discussão

A análise da literatura revela que, embora o campo da integração de acervos do patrimônio cultural tenha avançado significativamente em âmbito internacional, persistem lacunas relevantes tanto no plano conceitual quanto operacional, especialmente no contexto brasileiro. É evidente os avanços significativos no campo da integração de acervos culturais em âmbito internacional, ao mesmo tempo revela-se os desafios estruturais e conceituais que são persistentes no contexto brasileiro. A experiência de grandes iniciativas, como *Europeana*, DPLA, Sampo, CIDOC CRM e as redes baseadas em *Linked Open Data*, demonstra que a construção de ecossistemas digitais abertos não é um resultado automático da adoção de padrões tecnológicos, mas depende de uma conjugação de fatores institucionais, políticos, humanos e epistemológicos.

Doerr (2003, pp. 75-92), apontava para o grande desafio para a interoperabilidade semântica no campo do patrimônio cultural dizendo que, “a realidade da interoperabilidade semântica está se tornando frustrante. Só na área cultural, existem dezenas de padrões e centenas de estruturas de dados proprietárias, bem como centenas de sistemas de terminologia”.

No entanto, no plano internacional passados quase 1/4 de Século, observa-se a consolidação de modelos que priorizam a interoperabilidade semântica, a adoção de ontologias e a articulação entre diferentes padrões de metadados. Tais projetos revelam o potencial das tecnologias semânticas para transcender silos institucionais e permitir novas formas de descoberta, análise e reuso dos objetos culturais. Essa trajetória, contudo, não se deu sem resistência ou dificuldades, especialmente no que se refere à harmonização de dados legados, à adaptação de práticas institucionais e à necessidade de requalificação dos profissionais envolvidos. Fica claro que, por mais sofisticada que seja a infraestrutura tecnológica, sua eficácia só se concretiza quando alinhada à participação ativa dos profissionais da informação, à governança compartilhada e ao engajamento das comunidades.

Ao contrastar esse panorama com a realidade brasileira, emergem lacunas expressivas, especialmente quanto à capacidade de escalar iniciativas locais para políticas nacionais robustas. Ainda predomina a fragmentação institucional, a presença de sistemas legados e a falta de diretrizes unificadas para a publicação e integração de dados abertos. O país conta com experiências exitosas em instituições de referência, mas permanece dependente de projetos isolados e de soluções que raramente dialogam entre si. Soma-se a isso a ausência de investimento continuado em formação e atualização de equipes, a dificuldade em construir consensos interinstitucionais e a resistência à adoção de novos paradigmas de metadados, o que constitui obstáculos de difícil superação.

Esse cenário não é exclusivo do Brasil, mas reflete uma condição mais ampla do campo, marcada por grande diversidade de padrões, estruturas e vocabulários. Como observa Doerr (2003, pp. 75-92),

parece que a semântica por trás de um grande conjunto de estruturas (meta)dados diversas, provenientes de um domínio com muitas subdisciplinas, pode ser expressa por uma ontologia formal coerente baseada nas conceitualizações comuns dos especialistas do respectivo domínio, enquanto as próprias estruturas de entrada de dados frequentemente parecem resistir à fusão (tradução nossa).

Esse diagnóstico evidencia que, mesmo diante da possibilidade de harmonização conceitual via ontologias, as práticas e estruturas locais frequentemente resistem à integração efetiva. O'Neill & Stapleton (2022, pp. 891-903) ainda destacam que,

o processo de digitalização do patrimônio cultural exige a quebra de silos em diversas áreas — uma delas é a das disciplinas acadêmicas, para permitir a realização de um rico trabalho interdisciplinar. Isso estabelece as bases para a emancipação da segunda forma de silo, que são os silos de conhecimento, tanto tradicionais quanto digitais, mantidos em instituições individuais, como galerias, bibliotecas, arquivos e museus. A quebra de silos disciplinares é a chave para destravar esses silos institucionais de conhecimento. Equipes interdisciplinares, como desenvolvedores e bibliotecários, trabalham juntas para tornar os dados acessíveis como dados abertos na 'web semântica'.

Portanto, superar a fragmentação requer não apenas soluções tecnológicas ou modelos teóricos, mas um esforço coordenado de revisão institucional, colaboração intersetorial e investimento contínuo em capital humano.

A experiência internacional evidencia, assim, que a superação dessas barreiras passa, necessariamente, por iniciativas que promovam a colaboração interdisciplinar e a construção de ecossistemas abertos, integrando múltiplos atores, saberes e tecnologias.

Um dos pontos mais sensíveis identificados pela revisão é a necessidade de reconfigurar o papel do profissional da informação. Se, historicamente, as competências técnicas centradas em catalogação e preservação eram suficientes para garantir a gestão dos acervos, hoje é indispensável o

desenvolvimento de habilidades transversais, capacidade de liderança em projetos de integração, domínio de tecnologias semânticas e sensibilidade para mediação entre comunidades, sistemas e políticas. O protagonismo desses profissionais, aliado a uma formação continuada crítica e interdisciplinar, capaz de aliar domínio técnico em metadados, curadoria digital, Linked Data e ontologias à sensibilidade ética e cultural para mediação entre sistemas, comunidades e políticas públicas, é o que pode transformar a adoção de padrões internacionais em soluções genuinamente adaptadas às realidades e potencialidades nacionais.

A análise dos resultados revela que o avanço em direção a ecossistemas integrados e socialmente relevantes demanda, no contexto brasileiro, a superação de desafios que vão além do âmbito tecnológico. Embora iniciativas de digitalização tenham ampliado o acesso e a visibilidade dos acervos, tornam-se insuficientes se não forem acompanhadas de processos robustos de curadoria, avaliação de qualidade, interoperabilidade e efetiva participação comunitária. O estudo evidencia que o investimento em políticas de fomento ao capital humano, a formação continuada de equipes multidisciplinares e a criação de redes de colaboração institucional são elementos indispensáveis para romper a fragmentação e fortalecer os processos de integração.

Ademais, os resultados apontam que superar os silos institucionais constitui passo fundamental para a consolidação de ecossistemas digitais abertos, colaborativos e sustentáveis. A convergência observada na literatura e nas experiências internacionais indica que a combinação de esquemas híbridos de metadados com tecnologias de Linked Data, associada à adoção dos princípios FAIR e de identificadores persistentes, oferece caminhos concretos para o desenvolvimento de políticas nacionais de interoperabilidade e inovação no campo da memória. Para avançar nesse sentido, destaca-se a necessidade de compromisso institucional, colaboração intersetorial e investimento contínuo em pesquisa aplicada, formação profissional e experimentação de novos modelos de integração.

Por fim, destaca-se que a viabilidade de adaptação dos modelos internacionais ao contexto brasileiro — e, potencialmente, sul-americano — depende menos da simples importação de soluções tecnológicas e mais da capacidade de fomentar ecossistemas de inovação ancorados no protagonismo profissional, no fortalecimento de redes de cooperação e na construção de políticas institucionais abertas, inclusivas e colaborativas. Assim, o futuro da integração dos acervos culturais digitais passa, necessariamente, pelo reconhecimento do valor estratégico do capital humano e pela articulação de consensos em torno da abertura institucional e de dados, governança distribuída e responsabilidade social, capazes de promover uma memória coletiva mais aberta, conectada e sustentável.

5. Considerações Finais

A integração de acervos digitais do patrimônio cultural desponta como uma das tarefas mais estratégicas e urgentes para as instituições de memória, exigindo não apenas soluções tecnológicas, mas também políticas institucionais robustas, cooperação interinstitucional e, principalmente, o fortalecimento do capital humano. O exame crítico dos modelos internacionais evidencia que o êxito dessas experiências está diretamente relacionado à combinação entre inovação tecnológica, governança participativa, práticas colaborativas e investimento contínuo em formação profissional.

No contexto brasileiro, os desafios se aprofundam diante da fragmentação institucional, da ausência de políticas nacionais sustentáveis e da predominância de sistemas legados e projetos isolados. O enfrentamento desses entraves requer, sobretudo, a mobilização de redes de cooperação, a adoção de princípios abertos e o reconhecimento do papel estratégico do profissional da informação na liderança de processos de integração e curadoria digital.

A integração de arquivos, bibliotecas e museus — enquanto “lugares de memória” institucionalizados — transcende a dimensão meramente técnica e assume papel estratégico para a democracia, a cidadania informacional e a formação cultural das sociedades contemporâneas. Ao promover o acesso ampliado e integrado aos registros da memória coletiva, essas instituições ampliam sua função educativa, fortalecem o exercício dos direitos cidadãos e contribuem para a construção de identidades sociais plurais e críticas. Da mesma forma como enfatiza Alberch i Fugueras (n.d., tradução nossa) sobre, “a arquivística atual se caracteriza [...] como uma disciplina aliada da democracia ao constituir uma garantia do pleno exercício dos direitos cidadãos (acesso à informação, tomada de consciência da identidade, participação cidadã, transparência administrativa)”. Nesta mesma perspectiva, a superação de práticas fragmentadas potencializa os diversos mecanismos de lembrança, recuperação e preservação que caracterizam os arquivos, bibliotecas e museus enquanto lugares de memória adotados pelas sociedades para evitar o esquecimento e garantir a continuidade da identidade coletiva (Ramos & Miranda, 2021, pp. 68-85) .

Ou seja, a integração de acervos digitais potencializa impactos sociais e educativos concretos: amplia o acesso democrático à cultura e a memória, fomenta práticas inovadoras de ensino e pesquisa, fortalece a construção de identidades e apoia o exercício da cidadania informacional. Ao permitir o diálogo entre diferentes narrativas e fontes, contribui para sociedades mais críticas, plurais e engajadas.

Arquivos, Bibliotecas e Museus, mais do que concorrentes, essas instituições atuam de forma sinérgica para potencializar a difusão e o uso desses conjuntos informacionais, oferecendo as matérias-primas para a memória individual, social e coletiva (Ramos & Miranda, 2021, pp. 68-85). A quebra dos silos de dados e a integração de acervos, portanto, não apenas democratiza o acesso ao patrimônio, mas também amplia as possibilidades de apropriação crítica do conhecimento, promovendo uma educação cidadã, plural e voltada à inclusão social e à justiça informacional (cidadania informacional), consolidando as instituições de memória como pilares para sociedades mais participativas, diversas e informadas (Sena, 2023). Nesse sentido, a interoperabilidade se afirma como projeto social, educativo e político, essencial para a consolidação de uma sociedade mais informada, participativa e consciente de sua memória.

Por fim, recomenda-se que futuras agendas de pesquisa e políticas públicas priorizem, além da inovação tecnológica, o desenvolvimento de competências críticas, éticas e interdisciplinares, capazes de sustentar ecossistemas digitais verdadeiramente abertos, inclusivos e socialmente relevantes. Avançar rumo à integração efetiva dos acervos é investir, simultaneamente, em tecnologia, formação profissional e cidadania informacional, consolidando as instituições de memória e cultura como pilares do direito à informação, da educação e da democracia cultural.

Notas

¹ <https://www.europeana.eu/pt>

² <https://dp.la/>

³ <https://www.sotasampo.fi/en/>

⁴ <https://trove.nla.gov.au/>

⁵ <https://digitalnz.org/>

⁶ <https://www.w3.org/TR/rdf12-concepts/>

⁷ <https://www.w3.org/Addressing/>

⁸ <https://www.w3.org/DesignIssues/LinkedData.html>

⁹ <https://www.go-fair.org/fair-principles/>

10 <https://digital.bbm.usp.br/>
11 <https://tainacan.org/>
12 <https://bndigital.bn.gov.br/>
13 <https://www.brasilianaiconografica.art.br/>
14 <https://www.loc.gov/marc/bibliographic/>
15 [https://en.wikipedia.org/wiki/GLAM_\(cultural_heritage\)](https://en.wikipedia.org/wiki/GLAM_(cultural_heritage))
16 <https://pro.europeana.eu/page/edm-documentation>
17 <https://www.loc.gov/bibframe/>
18 <https://cidoc-crm.org/>
19 <https://www.loc.gov/ead/>
20 <https://www.dublincore.org/>
21 <https://www.iflstandards.info/fr/frbr/frbroo.html>
22 <https://en.wikipedia.org/wiki/API>
23 <https://seco.cs.aalto.fi/projects/ss/>
24 <https://www.w3.org/2004/02/skos/>
25 <https://www.w3.org/TR/owl2-primer/>
26 <https://www.getty.edu/research/tools/vocabularies/aat/>
27 <https://www.getty.edu/research/tools/vocabularies/tgn/>
28 <https://www.getty.edu/research/tools/vocabularies/ulan/>
29 <https://en.wikipedia.org/wiki/SPARQL>
30 <https://www.ica.org/resource/isadg-general-international-standard-archival-description-second-edition/>
31 <https://www.loc.gov/standards/vrcore/>
32 <https://www.w3.org/TR/xml-entity-names/>
33 <https://www.loc.gov/standards/mods/>
34 <https://www.loc.gov/standards/mets/mets-home.html>
35 <https://www.openarchives.org/pmh/>
36 <https://cidoc.mini.icom.museum/working-groups/lido/lido-overview/about-lido/what-is-lido/>
37 https://www.wikidata.org/wiki/Wikidata:Main_Page
38 <https://www.dbpedia.org/>
39 <https://www.geonames.org/>
40 <https://pro.dp.la/hubs/metadata-application-profile>
41 <https://repository.ifla.org/handle/20.500.14598/830>
42 International Federation of Library Associations and Institutions: <https://www.ifla.org/>
43 <https://repository.ifla.org/handle/20.500.14598/40.2>
44 <https://repository.ifla.org/handle/20.500.14598/3430>
45 <https://repository.ifla.org/handle/20.500.14598/835>
46 <https://icom.museum/en/committee/international-committee-for-documentation/>
47 <https://viaf.org/en>
48 <https://www.w3.org/TR/rdf12-turtle/>
49 <https://json-ld.org/>

Referências

Alberch i Fugueras, R. (n.d.). Archivos, calidad, memoria y derechos democráticos. Los desafíos del siglo XXI. <https://www.marilia.unesp.br/Home/Extensao/CEDHUM/texto02.pdf>

Arakaki, A. C. S. (2020). O modelo IFLA Library Reference Model e o Linked Data. *Informação & Informação*, 25(3), 163–186. <https://doi.org/10.5433/1981-8920.2020v25n3p163>

Bettencourt, A. M., & Marcondes, C. H. (2019). Elementos para uma política brasileira de acesso integrado, utilização e preservação de acervos digitais em memória e cultura. *Pragmatizes: Revista Latino-Americana de Estudos em Cultura*, 9(16), 44–61. <http://periodicos.uff.br/pragmatizes/%20article/view/27518/16801>

Buckland, M. K. (2006). Descrição e pesquisa: metadados como infra-estrutura. *Brazilian Journal of Information Science: research trends*, 0, 3–14. <https://doi.org/10.36311/1981-1640.2006.v0n0.02.p3>

Charles, V., & Isaac, A. (2015). *Enhancing the Europeana Data Model (EDM): White Paper*. Europeana Foundation. 1-21. https://pro.europeana.eu/files/Europeana_Professional/Publications/EDM_WhitePaper_17062015.pdf

Concordia, C., Gradmann, S., & Siebinga, S. (2010). Not just another portal, not just another digital library: A portrait of Europeana as an application program interface. *IFLA Journal*, 36(1), 61–69. <https://doi.org/10.1177/0340035209360764>

Digital Public Library of America (DPLA). (2017). Introduction to the DPLA Metadata Application Profile, version 5.0. <https://pro.dp.la/hubs/metadata-application-profile>

Doerr, M. (2003). The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine*, 24(3), 75-92. <https://doi.org/10.1609/aimag.v24i3.1720>

Europeana Foundation. (n.d.). *EDM Factsheet*. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Factsheet.pdf

Gaitanou, P., Andreou, I., Sicilia, M.-A., & Garoufallou, E. (2024). Linked data for libraries: Creating a global knowledge space, a systematic literature review. *Journal of Information Science*, 50(1), 204–244. <https://doi.org/10.1177/01655515221084645>

Hermon, S., & Niccolucci, F. (2021). FAIR Data and Cultural Heritage Special Issue Editorial Note. *International Journal on Digital Libraries*, 22, 251–255. <https://doi.org/10.1007/s00799-021-00309-8>

Hyvönen, E. (2023). Digital humanities on the Semantic Web: Sampo model and portal series. *Semantic Web*, 14(4), 729–744. <https://doi.org/10.3233/SW-223034>

Isaac, A. (2013). *Europeana Data Model Primer*. Europeana Foundation. 1-35. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf

International Organization for Standardization. (2023). ISO 21127:2023 – Information and documentation: A reference ontology for the interchange of cultural heritage information (3rd ed.). Geneva: ISO.

Kaldeli, E. (2023). *Combining AI tools with human validation to enrich cultural heritage metadata*. Europeana Pro. <https://pro.europeana.eu/post/combining-ai-tools-with-human-validation-to-enrich-cultural-heritage-metadata>

Lemos, D. L. da S., Coelho Junior, A., & Martins, D. L. (2023). Modelos de diagnóstico de qualidade de dados no domínio do patrimônio cultural: uma revisão sistemática de literatura. *Perspectivas em Ciência da Informação*, 28, e46064, 1-26. <https://doi.org/10.1590/1981-5344/46064>

Marcondes, C. H. (2016). Metadados e a Web Semântica: Oportunidades e desafios para integração de dados, objetos digitais e coleções digitais em ambientes digitais na Web. *Perspectivas em Ciência da Informação*, 21(2), 61–83. <https://doi.org/10.1590/1981-5344/2735>

Marcondes, C. H. (2019). Uma classificação de tipos de objetos de patrimônio para integração de acervos digitais. *Scire*, 25(2), (pp.45-52).

Marcondes, C. H. (2022). Um modelo de curadoria para integrar acervos digitais em Memória e Cultura publicados na Web como Dados Abertos Interligados. *Frontiers of Knowledge Representation (FRC: Front. Repr. Conh.)*, 2(2), 1-31. <https://doi.org/10.5281/zenodo.8352124>

Marcondes, C. H. (2023). A curatorial model for digital heritage collections as linked open data and named graphs. *Scire*, 29(2), 85–95. <https://doi.org/10.54886/scire.v29i2.4898>

O'Neill, B., & Stapleton, L. (2022). Digital cultural heritage standards: From silo to semantic web. *AI & Society*, 37, 891–903. <https://doi.org/10.1007/s00146-021-01371-1>

Ramos, T. O., & Miranda, Z. D. de. (2023). O inter-relacionamento entre documentos de Arquivo, Biblioteca e Museu:: Memorial - um sistema em definição. *Revista Fontes Documentais*, 4(1), 68–85. <https://periodicos.ufba.br/index.php/RFD/article/view/57691>

Ribeiro, A. T. M., Ribas, R. A., & Amaral, S. (2020). Integração e interoperabilidade entre acervos arquivísticos, bibliográficos e museológicos: estudo exploratório de plataformas digitais brasileiras. *Biblos: Revista do Instituto de Ciências Humanas e da Informação*, 34(2), 247–264. <https://doi.org/10.14295/biblos.v34i2.11864>

Sacramento, C., Baião, F. A., & Ferreira, S. B. L. (2015). Um estudo sobre metadados para preservação e integração de acervos digitais heterogêneos. *Relatórios Técnicos do Departamento de Informática Aplicada da UNIRIO*, (0002/2015), 1–49.

Salse, M., Guallar-Delgado, J., Jornet-Benito, N., Mateo Bretos, M. P., & Silvestre-Canut, J. O. (2024). GLAM metadata in museums and university collections: A state-of-the-art (Spain and other European countries). *Global Knowledge, Memory and Communication*, 73(4/5), 477–495. <https://doi.org/10.1108/GKMC-06-2022-0133>

Sena, P. M. B.. (2023). Justiça Informacional Em Ciência, Tecnologia E Inovação No Brasil: Reflexões E Ações Necessárias Em Ciência Da Informação. *Encontros Bibli*, 28(Spe), E93046. <https://doi.org/10.5007/1518-2924.2023.E93046>

Silva, A. M. da. (2013). A ciência da informação e a transição paradigmática. In M. A. Rendón Rojas (Coord.), *El objeto de estudio de la bibliotecología/documentación/ciencia de la información: propuestas, discusión, análisis y elementos comunes* (pp. 85-103). México: UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información.

Silva, E. L., & Terra, A. L. (2024). Semantic interoperability in digital cultural heritage: The Europeana Data Model and the semantic web. *International Federation of Library Associations and Institutions*, 50(1), 93–107. <https://doi.org/10.1177/03400352231202506>

Siqueira, J., & Martins, D. L. (2022). Workflow models for aggregating cultural heritage data on the web: A systematic literature review. *Journal of the Association for Information Science and Technology*, 73, 204–224. <https://doi.org/10.1002/asi.24498>

Vital, L. P., & Brascher, M. (2017). Modelo conceitual na descrição arquivística: Uma análise sobre a representação temática. In F. A. Pinho & J. A. C. Guimarães (Eds.), *Memória, tecnologia e cultura na organização do conhecimento* (pp. 202–207). Editora UFPE.

Wijesundara, C., & Sugimoto, S. (2018). Metadata Model for Organizing Digital Archives of Tangible and Intangible Cultural Heritage and Linking Cultural Heritage Information in Digital Space. *LIBRES: Library and Information Science Research Electronic Journal*, 28(2), 58–80. Disponível em: <https://www.libres-ejournal.info/2706/>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, Article 160018. <https://doi.org/10.1038/sdata.2016.18>