

Eixo 3: Prática em Classificação, Catalogação e Indexação

## O MODELO LRMoo E AS PERSPECTIVAS DE INTEGRAÇÃO DE DADOS DE BIBLIOTECAS E MUSEUS

### *The LRMoo model and the perspectives of libraries and museums' data integration*

**Isabel Cristina Ayres da Silva Maringelli**

Doutorado - Fundação Escola de Sociologia e Política de São Paulo (FESPSP), São Paulo, SP, Brasil

ORCID: <https://orcid.org/0000-0002-8861-7421>

**José Fernando Modesto da Silva**

Pós-Doutorado - Escola de Comunicações e Artes da Universidade de São Paulo (ECA/USP), São Paulo, SP, Brasil

ORCID: <https://orcid.org/0000-0002-0535-5471>

#### Resumo

Este estudo explora o modelo Library Reference Model Object Oriented (LRMoo) e suas possibilidades de integração de dados de bibliotecas e museus, destacando a criação de representações gráficas baseadas nas relações entre uma obra de arte e suas reproduções. A pesquisa é qualitativa e descritiva-exploratória, utilizando procedimentos metodológicos que envolvem a conversão de dados para o formato RDF e a utilização de vocabulários controlados para evitar conflitos semânticos. O estudo escolheu a obra "O Caipira Picando Fumo" de Almeida Júnior para ilustrar o mapeamento com o CIDOC CRM e o LRMoo. Os resultados indicam que a integração de dados bibliográficos e museológicos pode ser efetivamente analisada e modelada a partir da perspectiva do LRMoo, o que abre novas possibilidades para a colaboração entre instituições culturais e a construção de uma base de

conhecimento acessível e socialmente transmissível. A pesquisa conclui que, embora preliminares, os objetivos foram alcançados, permitindo o planejamento de estudos mais aprofundados e a inclusão de dados arquivísticos para enriquecer os recursos informacionais).

**Palavras-Chave:** Integração de dados; LRMoo; CIDOC CRM; Bibliotecas; Museus.

### **Abstract**

This study explores the Library Reference Model Object Oriented (LRMoo) and its possibilities for integrating library and museum data, highlighting the creation of graphical representations based on the relationships between a work of art and its reproductions. The research is qualitative and descriptive-exploratory, using methodological procedures that involve converting data to RDF format and using controlled vocabularies to avoid semantic conflicts. The study chose the work "O Caipira Picando Fumo" by Almeida Júnior to illustrate mapping with CIDOC CRM and LRMoo. The results indicate that the integration of bibliographic and museological data can be effectively analyzed and modeled from the perspective of LRMoo, which opens new possibilities for collaboration between cultural institutions and the construction of an accessible and socially transmissible knowledge base. The research concludes that, although preliminary, the objectives were achieved, allowing for the planning of more in-depth studies and the inclusion of archival data to enrich informational resources.

**Keywords:** Data integration; LRMoo; CIDOC CRM; Libraries; Museums.

## **1 INTRODUÇÃO**

A tecnologia muda continuamente e a web, da mesma forma, continua a evoluir. Neste contínuo a web se torna mais personalizada e personalizável. Computadores e algoritmos nos ajudam a tomar decisões e classificar informações com mais precisão. A web também tem se tornado mais inteligente à medida que se desenvolvem modelos semânticos para estruturar dados visando melhor

recuperação por mecanismos de busca ou sua integração entre bases e sistemas de armazenamento e pesquisa (Fay; Seuers, 2012).

A Web Semântica tem facilitado a integração de dados sobre o patrimônio cultural, composto por coleções de naturezas diversas, e necessidades e processos distintos de organização. Esses dados refletem a complexidade do domínio do patrimônio cultural, “caracterizado por um panorama de dados altamente heterogêneos que abrange uma variedade de tipos de recursos” (Lodi et al., 2017, p. 11).

Embora o modelo LRMoo (*Library Reference Model Object Oriented*) seja amplamente divulgado por instituições como o *International Council of Museums* (ICOM) e a *International Federation of Library Associations and Institutions* (IFLA), além de ser abordado em estudos de caso apresentados em eventos dessas entidades, ainda existem lacunas significativas na literatura nacional a respeito desse modelo. A principal questão a ser explorada é a viabilidade do LRMoo para a integração de dados no âmbito do patrimônio cultural. Este estudo tem como objetivo investigar as possibilidades de aplicação do LRMoo, com foco na criação de representações gráficas que ilustram as relações entre obras de arte e suas reproduções.

## 2 SISTEMA PERGAMUM E O RDA

Segundo Gil (2007), a pesquisa exploratória visa a compreensão de um problema de forma a torná-lo mais explícito. Quanto a abordagem qualitativa, Santade (2014), destaca o aspecto de estudo não estatístico, que identifica e analisa dados de difícil mensuração, mas com base em análises realizadas por um pesquisador-observador. Neste contexto, esta pesquisa é de natureza qualitativa e do tipo descritiva-exploratória, pois explora as possibilidades de emprego do modelo LRMoo para integração de dados de bibliotecas e museus por meio de pesquisa bibliográfica e documental.

Saliente-se, que a elaboração do mapeamento ilustrado, na Figura 3, baseou-se no gráfico do modelo LRMoo, e os dados foram coletados no catálogo online da Biblioteca de Artes Visuais da Pinacoteca de São Paulo e no acervo museológico.

### 3 DESENVOLVIMENTO

Para que o diagrama modelado em *Functional Requirements for Bibliographic Records – object-oriented* (FRBRoo), compreensível por humanos, seja legível por máquina por meio da utilização do *Resource Description Framework* (RDF), é necessário realizar uma conversão (Chen; Ke, 2013). Diversos instrumentos podem ser utilizados para essa tarefa, dentre os quais citamos o TopBraid, um “software para edição de ontologias utilizado para criação de ontologias para a Web Semântica e construção de aplicações semânticas” (Chen; Ke, 2013, p. 630) e o Protégé “software de código aberto desenvolvido pela Universidade de Stanford” (Lima, 2008, p. 142).

No caso de dados heterogêneos, outras etapas podem ser necessárias antes desse processo, como, por exemplo, a conversão de dados em formato *Machine Readable Cataloging* (MARC21) de uma base de dados bibliográfica. O *World Wide Web Consortium* (W3C) recomenda algumas ferramentas para essa conversão, além do marc2rdf, por exemplo. Ressalta-se a importância dos vocabulários controlados nesse processo, pois no momento da conversão, o software pode fazer o cruzamento de conceitos, adotando o que for mais apropriado, evitando que diferenças de idiomas gerem conflitos (Lisena et al., 2018, p. 198):

Dado que uma meta compartilhada da comunidade de patrimônio cultural é a melhoria do acesso às artes visuais e à informação de cultura material, os vocabulários controlados são essenciais. Eles são necessários durante a fase de indexação porque, sem eles, os catalogadores não utilizarão consistentemente o mesmo termo para referir-se à mesma pessoa, lugar ou coisa (Harpring, 2016, p. 37)

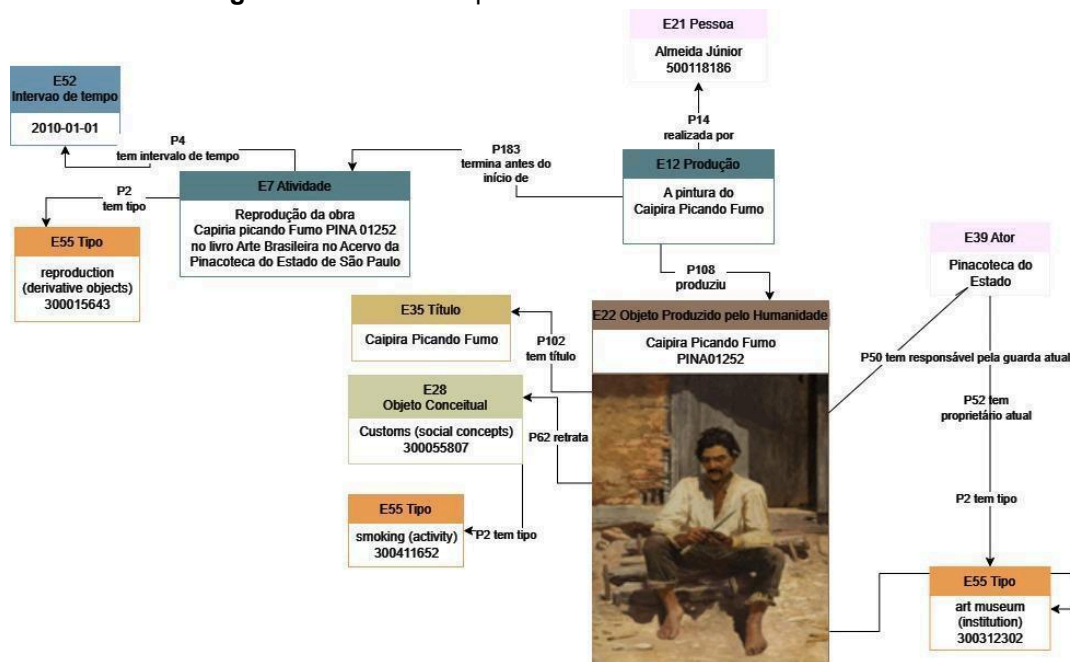
Para a criação de representações gráficas com o CIDOC Conceptual Reference Model (CRM) e com o LRMoo escolhemos a obra “O Caipira Picando

Fumo”, de autoria de Almeida Júnior. Sua presença no imaginário é constante, sendo uma das mais conhecidas obras produzidas pelo artista, retratando o homem rural, isolado das relações sociais urbanas. O picador de fumo, mantendo sua postura concentrada expondo sua faca de forma contundente, deixando-a entre si mesmo e o espectador, e dessa forma protege-se e protege seu lugar frágil no mundo pela violência possível (Coli, 2002).

### 3.1 MAPEAMENTO COM CIDOC-ROM

A figura 1 demonstra o mapeamento das relações hierárquicas e os principais relacionamentos propostos pelo CIDOC CRM. O diagrama foi elaborado em *Unified Modeling Language* (UML), utilizada para modelar softwares baseados no paradigma de orientação a objetos (Guedes, 2011, p. 19).

Figura 1 - Gráfico “Caipira Picando Fumo” - CIDOC CRM



Fonte: Elaborado pelos autores (2023)

Na Figura 1, a atividade central é a reprodução da pintura em um catálogo. Conforme indicado pelo CIDOC CRM, “as instâncias das classes são representadas por rótulos informativos em vez de identificadores” (International Federation of

Library and Institutions, 2023a, p. 37). As propriedades são representadas como setas e sempre que possível, incluiu-se o código correspondente ao vocabulário utilizado, conforme explicitado no Quadro 1.

**Quadro 1** - Vocabulários utilizados na obra “Caipira Picando Fumo”

Classe	Vocabulário	Código
E22 Human-Made Object	Pinacoteca	Caipira Picando Fumo –Tombo PINA01252
E28 Conceptual Object	Getty AAT	Customs (social concepts) 300055807
E55 Type	Getty AAT	Reproduction (derivative objects) 300015643
E55 Type	Getty AAT	smoking (activity) 300411652
E55 Type	Getty AAT	art museum (institution) 300312302
E21 Person	Getty ULAN	Almeida Júnior 500118186

Fonte: Elaborado pelos autores.

Após a criação do gráfico em CIDOC CRM, foi gerado um arquivo em RDF, com a utilização de uma ferramenta disponível na página do CIDOC Services, conforme demonstrado na Figura 2:

**Figura 2** – Detalhe do arquivo em RDF da obra “Caipira Picando Fumo” (detalhe)

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <rdf:RDF xmlns:crm="http://www.cidoc-crm.org/cidoc-crm/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3 <rdf:Description rdf:about="http://www.cidoc-crm-from-draw.io/"
4 <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
5 <rdfs:label>CIDOC-CRM v7.1.1.2 example produced by input file: 'Teste1.drawio (4).xml'</rdfs:label>
6 <rdfs:comment>
7 * INFO: The diagram from file: 'Teste1.drawio (4).xml' specified: 11 instances connected with: 11 properties.
8 </rdfs:comment>
9 </rdf:Description>
10 <rdf:Description rdf:about="http://www.cidoc-crm-from-draw.io/yZzM00kknx0OFva2ydp8-10">
11 <rdf:type rdf:resource="http://www.cidoc-crm.org/cidoc-crm/E22_Human-Made_Object"/>
12 <rdfs:label>Caipira Picando Fumo PINA01252</rdfs:label>
13 <crm:P102_has_title rdf:resource="http://www.cidoc-crm-from-draw.io/yZzM00kknx0OFva2ydp8-7"/>
14 <crm:P138i_has_representation rdf:resource="http://www.cidoc-crm-from-draw.io/yZzM00kknx0OFva2ydp8-12"/>
15 <crm:P50_has_current_keeper rdf:resource="http://www.cidoc-crm-from-draw.io/HW-IxbHXULawjK4Wl_4i-3"/>
16 <crm:P52_has_current_owner rdf:resource="http://www.cidoc-crm-from-draw.io/HW-IxbHXULawjK4Wl_4i-3"/>
17 <crm:P62_depicts rdf:resource="http://www.cidoc-crm-from-draw.io/yZzM00kknx0OFva2ydp8-54"/>
18 </rdf:Description>
  
```

Fonte: Elaborado pelos autores baseados em Drawio ([2023]).

A conversão do gráfico para o formato RDF permite que o arquivo seja legível por máquina e possa ser compartilhado na Web.

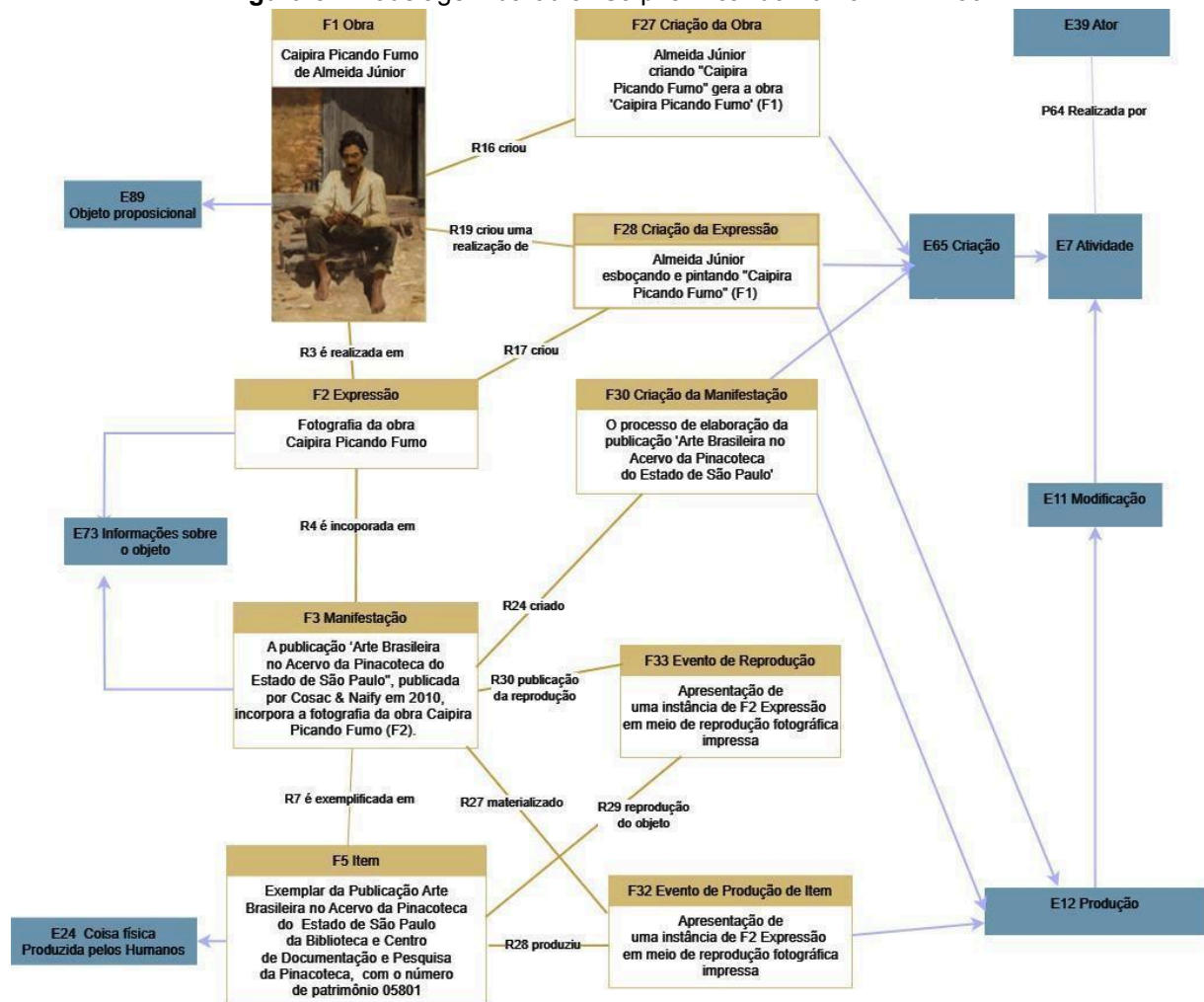
### 3.2 MAPEAMENTO COM LRMoo

A título de ensaio, foi elaborada uma representação semântica com o LRMoo relacionando-se a obra “Caipira Picando Fumo” e a publicação de uma reprodução dela por meio de uma fotografia publicada em um livro de arte (Figura 3).

As conexões entre as classes dos modelos se dão a partir das classes superiores do CIDOC CRM (indicadas em azul). O gráfico inclui as classes F1 Obra,

F2 Expressão, F3 Manifestação e F5 Item. Tais classes são derivadas do núcleo composto pelas entidades *Work*, *Expression*, *Manifestation* e *Item* (WEMI) do *Functional Requirements for Bibliographic Records* (FRBR) subordinadas às classes do CIDOC CRM (indicadas em azul). As classes F27 Criação de Obra, F28 Criação da Expressão, e F30 Criação da Manifestação são subclasses das classes CIDOC CRM E65 Criação ou E12 Produção, que estão subordinadas à E7 Atividade que é realizada por E39 Ator. Os relacionamentos são indicados pelas propriedades iniciadas pela letra P ou pela letra R.

Figura 3 - Modelagem da obra “Caipira Picando Fumo” – LRMoo



Fonte: Elaborado pelos autores baseados em International Federation of Library and Institutions (2023a).

Os registros bibliográficos são elaborados de acordo com o *Anglo American Cataloging Rules Second Edition* (AACR2), não estando, portanto, no padrão

recomendado para dados abertos, nem modelados conforme o IFLA LRM ou o *Resource Description and Access* (RDA). A Figura 4 contém os metadados bibliográficos da publicação que inclui a reprodução da obra publicada em um livro (formato MARC21 bibliográfico):

**Figura 4** - Registro bibliográfico da publicação “Arte brasileira” (detalhe)

245 @00^aArte brasileira na Pinacoteca do Estado de São Paulo : ^bdo século XIX aos anos 1940 / ^corganização Taisa Helena P. Palhares; textos de Annateresa Fabris...[et al.]

260 @##^aSão Paulo : ^bCosac Naify, ^c2009.

300 @##^a237 p. : ^bil.

500 @##^aÍndice de artistas e obras: p. 234.

500 @##^aBiografias: p. 209.

505 @0#^aArtistas: Victor Meirelles, Niclas-Antoine Taunay, Jean Baptiste Debret, Ângelo Agostini, Pedro Américo, Jorge Grimm, João Batista Castagneto, Antonio Parreiras, Henrique Bernardelli, Eugène Delacroix, Robert Trew, Thomas Lawrence, Rodolfo Bernardelli, Alexandre Cabanel, John Everett Millais Almeida Júnior, Antonio Ferrigno, Modesto Brocos, Claude Monet, Pedro Alexandrino, Eliseu Visconti, Anita Malfatti, Victor Brecheret, Tarsila do Amaral, Lasar Segall, Flávio de Carvalho, Candido Portinari,

**Fonte:** Pinacoteca de São Paulo (2023).

Até o momento de elaboração deste texto, os dados do acervo museológico não seguiam um padrão de metadados específico. O estabelecimento de diretrizes para registro dos dados foi realizado internamente por meio da publicação de um manual de catalogação.

A proposta do LRMoo alinha-se aos princípios da Web Semântica, especialmente no que se refere ao compartilhamento e reuso de dados. A integração de acervos culturais de diferentes naturezas é um desafio constante tanto para os profissionais da área quanto para os usuários desses acervos.

## 4 RESULTADOS

A relação das classes oriundas do núcleo WEMI com as classes ditas de “criação” nos permitem contextualizar o processo que precede a elaboração da manifestação (F3). Podemos questionar o quanto a reprodução de uma pintura por meios fotográficos, transformada em uma imagem impressa em um livro, pode ser considerada como expressão da obra original. Embora a pintura possua uma bidimensionalidade que é perdida na fotografia e no livro, o processo de criação da manifestação (F30) e do evento de reprodução (F33) podem auxiliar na

contextualização entre o conteúdo da manifestação (F3) e o meio e condições em que foi produzida.

A modelagem realizada neste estudo exemplifica a complexidade das relações documentais existentes entre uma obra de arte e sua reprodução, as quais nem sempre são explicitadas nos sistemas de informação disponibilizados para o usuário, ou seja, depreende-se que a busca por informações sobre determinado artista ou obra pode exigir a consulta a bancos de dados diversos que utilizam vocabulários diferentes, mas que em realidade referem-se aos mesmos objetos e atores. O LRMoo propõe a elaboração de dados semânticos e enriquecidos, favorecendo a integração entre os dados dos acervos.

Conclui-se que a integração de dados bibliográficos e dados museológicos pode ser abordada a partir da perspectiva do LRMoo. Sendo uma extensão do CIDOC CRM, utilizado em larga escala pela comunidade museológica, somados aos elementos do IFLA LRM, ele traz perspectivas promissoras no campo da modelagem.

## 5 CONSIDERAÇÕES FINAIS

Para que nossa proposta possa ser testada e validada será realizada a conversão dos dados reais para o RDF, com a utilização dos knowledge graphs, também conhecidos como "gráficos de conhecimento", os quais são criados a partir de fontes de dados estruturadas e abertas e funcionam (Coneglian et al., 2017).

Serão incluídos também dados do acervo arquivístico, de forma a completar os recursos informacionais relacionados à obra de arte aqui estudada.

Ainda que os resultados apresentados sejam de caráter preliminar, considera-se que os objetivos foram alcançados, pois a partir do ensaio realizado é possível planejar estudos mais aprofundados, com metadados do mundo real.

Também se considera que o aumento da colaboração entre instituições culturais, proporcionado por modelos que dialoguem, permite construir ligações fortes para uma maior eficácia na comunicação com o público.

## 6 REFERÊNCIAS

- CHEN, Y. N.; KE, H. R. FRBRoo-based approach to heterogeneous metadata integration. **Journal of Documentation**, London, v. 69, n. 5, p. 623-637, 2013. DOI: 10.1108/JD-07-2012-0086.
- COLI, J. A violência e o caipira. **Estudos Históricos**, Rio de Janeiro, n. 30, p. 23-31, 2002. Disponível em: <https://periodicos.fgv.br/reh/article/view/2172/1311>. Acesso em: 30 abr. 2023.
- CONEGLIAN, C. S. et al. A experiência do usuário nos mecanismos de busca knowledge graph e o knowledge vault. **Informação@Profissões**, Londrina, v. 6, n. 2, p. 35-59, jul./dez. 2017. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/infoprof/article/view/33103/23614>. Acesso em: 10 dez. 2023.
- DRAWIO. [2023]. Disponível em: [https://isl.ics.forth.gr/cidoc\\_services/drawioXMLToTriples](https://isl.ics.forth.gr/cidoc_services/drawioXMLToTriples). Acesso em: 20 dez. 2023.
- FAY, R. M.; SEUERS, M. P. **Semantic Web Technologies and social searching for librarians**. Chicago: ALA, 2012.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2007.
- GUEDES, G. T. A. **UML 2: uma abordagem prática**. São Paulo: Novatec Editora, 2011.
- HARPRING, P. **Introdução aos vocabulários controlados: terminologia para arte, arquitetura e outras obras culturais**. São Paulo: Pinacoteca do Estado, 2016.
- INTERNATIONAL FEDERATION OF LIBRARY AND INSTITUTIONS. LRMoo Working Group with the CIDOC CRM special interest group. **LRMoo object-oriented definition and mapping from IFLA LRM**. [S. l.]: IFLA, 2023a. Disponível em: <https://cidoc-crm.org/frbroo/ModelVersion/version-0.9.6>. Acesso em: 10 jul. 2024.
- INTERNATIONAL FEDERATION OF LIBRARY AND INSTITUTIONS. **Working Group on FRBR – CIDOC CRM Harmonization – IFLA**. Haia: IFLA, 2023b. Disponível em: <https://www.ifla.org/g/bcm-rg/working-group-on-frbr-cidoc-crm-harmonization>. Acesso em: 16 jul. 2024.
- LIMA, J. A. de O. **Modelo genérico de relacionamentos na organização da informação jurídica e legislativa**. 2008. Tese (Doutorado em Ciência da

Informação) – Universidade de Brasília, Brasília, 2008. Disponível em: <http://repositorio.unb.br/jspui/handle/10482/3398>. Acesso em: 10 jul. 2024.

LISENA, P. et al. Improving (Re-) Usability of Musical Datasets: An Overview of the DOREMUS Project. **Bibliothek Forschung und Praxis**, Berlin, v. 42, n. 2, p. 194-205, June 2018.

LODI, G. et al. Semantic Web for Cultural Heritage Valorisation. In: HAI-JEW, S. (ed.). **Data Analytics in Digital Humanities**. Multimedia Systems and Applications. Springer, Cham, 2017.

PINACOTECA DE SÃO PAULO. **Banco de dados do acervo bibliográfico**. 2023. Disponível em: <http://biblioteca.pinacoteca.org.br:9090/>. Acesso em: 30 ago. 2024.

SANTADE, M. S. B. a metodologia de pesquisa: instrumentais e modos de abordagem. In: SIMÕES, Darcilia M. P.; GARCÍA, Flavio. A pesquisa científica como linguagem e práxis. Rio de Janeiro: Dialogarts, 2014.