

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

Ana Carolina Novaes de Mendonça

**Metadado de gênero em registros de autoridade: proposta de diretrizes e
enriquecimento semântico**

São Carlos/SP

2025

Ana Carolina Novaes de Mendonça

Metadado de gênero em registros de autoridade: proposta de diretrizes e enriquecimento semântico

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade da Universidade Federal de São Carlos (UFSCar), como parte das exigências para obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Conhecimento, Tecnologia e Inovação

Linha de pesquisa: Tecnologia, Informação e Representação

Orientadora: Profa. Dra. Ana Carolina Simionato Arakaki



Dados e Metadados

grupo de pesquisa e extensão

São Carlos/SP

2025

Mendonça, Ana Carolina Novaes de

Metadado de gênero em registros de autoridade:
proposta de diretrizes e enriquecimento semântico / Ana
Carolina Novaes de Mendonça -- 2025.
135f.

Dissertação (Mestrado) - Universidade Federal de São
Carlos, campus São Carlos, São Carlos
Orientador (a): Ana Carolina Simionato Arakaki
Banca Examinadora: Fabiano Ferreira de Castro, Rachel
Cristina Vesu Alves
Bibliografia

1. Enriquecimento semântico. 2. Gênero. 3. Metadados.
I. Mendonça, Ana Carolina Novaes de. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Arildo Martins - CRB/8 7180



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Educação e Ciências Humanas
Programa de Pós-Graduação em Ciência da Informação

Folha de Aprovação

Defesa de Dissertação de Mestrado da candidata Ana Carolina Novaes de Mendonça, realizada em 15/04/2025.

Comissão Julgadora:

Profa. Dra. Ana Carolina Simionato Arakaki (UFSCar)

Prof. Dr. Fabiano Ferreira de Castro (UFSCar)

Profa. Dra. Rachel Cristina Vesu Alves (UNESP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Informação.

“Dizem que a vida é para quem sabe viver,
mas ninguém nasce pronto. A vida é para
quem é corajoso o suficiente para se
arriscar e humilde o bastante para
aprender.”

Clarice Lispector

AGRADECIMENTOS

À minha mãe, Maria José, meu maior exemplo de força e dedicação.

Ao meu pai, Fábio (in memoriam), cuja presença e ensinamentos continuam vivos em mim.

Ao meu irmão, Pedro, que esteve ao meu lado desde o primeiro momento.

Aos meus tios, Beth, Carlos, Regina e Luiz pelo apoio e compreensão.

Ao meu padrasto Attilio pelo incentivo.

Aos que se foram durante essa jornada, vô Pedro e vô Edson. À toda a minha família, que mesmo à distância, sempre me acompanhou, torceu e apoiou.

Aos meus amigos que são família, e àqueles que também me emprestaram as suas, em especial, Bianca, Carol, Camila, Chiara, Duda Casado, Duda Wurmbauer, Fel, Fran, Gui, João, Juju, Luana, Pedro, Rachel, Tati, Victoria, Vini, Vivi, Sarah e Yumi.

À Ju, que me acolheu em São Carlos.

À Dayane, colega de pesquisa e que se tornou uma grande amiga ao longo dessa caminhada.

À minha orientadora, professora Ana Carolina Simionato Arakaki, por sua confiança, orientação e paciência, e por me ensinar tanto sobre resiliência.

Ao professor Felipe Augusto Arakaki, pelo empurrãozinho que me levou à vida acadêmica e pela confiança depositada em mim. Serei eternamente grata, profs.

À turma 8 do PPGCI/UFSCar, pelas trocas dentro e fora da sala de aula, pelos almoços, cafés e bolos, os dias que tínhamos aula eram mais leves e felizes. Um carinho especial para ao César, Gabi, Luana, Maísa e Vitória, que foram parte essencial dessa trajetória.

Aos professores Fabiano e Rachel pelas contribuições valiosas ao trabalho.

Aos professores e funcionários do PPGCI, por todo o conhecimento compartilhado, pelas aulas enriquecedoras e pelo incentivo ao longo dessa caminhada.

Ao Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) pela oportunidade de ser pesquisadora.

Aos profissionais das bibliotecas nacionais que responderam à consulta.

A todos os que direta ou indiretamente contribuíram para a realização desta pesquisa.

RESUMO

A presença do metadado de gênero em registros de autoridade tem se tornado um tema emergente para a área de Ciência da Informação, especialmente diante da crescente demanda por sistemas de organização da informação e do conhecimento mais inclusivos, interoperáveis e sensíveis à diversidade identitária. Em um contexto marcado pela transformação das práticas catalográficas, impulsionadas por modelos conceituais, pelos avanços das Tecnologias da Informação e Comunicação e pela adoção de padrões de metadados estruturados, surge a necessidade de repensar a forma como atributos como o gênero são representados em catálogos bibliográficos. Esta pesquisa teve como objetivo analisar a presença e o tratamento do metadado de gênero em registros de autoridade mantidos por bibliotecas nacionais, propondo diretrizes e estratégias baseadas em enriquecimento semântico para qualificar essa representação nos sistemas informacionais. Trata-se de uma investigação de natureza qualitativa, com abordagem descritiva e exploratória, fundamentada em levantamento bibliográfico, análise documental e observação direta de catálogos públicos. Foram analisadas treze bibliotecas nacionais, selecionadas por adotarem diferentes padrões de metadados, buscando identificar os campos utilizados para representar o gênero, os valores atribuídos, os vocabulários controlados empregados, bem como a eventual vinculação a identificadores externos, como *Wikidata* e *VIAF*. A pesquisa também examinou diretrizes institucionais, documentos técnicos e políticas públicas relacionadas à inclusão do metadado de gênero, quando disponíveis. Os resultados indicam que, embora os padrões de metadados já prevejam campos específicos para o registro do gênero, sua aplicação é ainda limitada, majoritariamente binária e desprovida de padronização entre as instituições analisadas. Observou-se a ausência de políticas institucionais claras para o uso do campo, bem como a predominância de práticas descritivas que negligenciam a autodeclaração e a fluidez das identidades de gênero. Por outro lado, foram identificadas experiências positivas que envolvem o uso de vocabulários controlados sensíveis à diversidade, vínculos com URIs externas e registros alinhados a ontologias abertas. Conclui-se que a qualificação do metadado de gênero nos registros de autoridade, por meio do uso de tecnologias de dados conectados e do enriquecimento semântico, pode fortalecer a expressividade, a interoperabilidade e a justiça informacional nos catálogos bibliográficos. A adoção de diretrizes técnicas e éticas para essa representação é essencial para assegurar a visibilidade de autoras, pessoas trans, não binárias e outras identidades frequentemente sub-representadas nos sistemas tradicionais. A pesquisa evidencia, assim, a importância de avanços normativos e políticos para promover a diversidade e a equidade nos ambientes informacionais.

Palavras-chave: Enriquecimento semântico; Gênero; Metadados; Registros de autoridade; *Web Semântica*.

ABSTRACT

The presence of gender metadata in authority records has become an emerging topic in the field of Information Science, especially in light of the growing demand for more inclusive, interoperable, and identity-aware systems of information and knowledge organization. In a context marked by the transformation of cataloging practices, driven by conceptual models, advances in Information and Communication Technologies, and the adoption of structured metadata standards, there arises a need to rethink how attributes such as gender are represented in bibliographic catalogs. This research aimed to analyze the presence and treatment of gender metadata in authority records maintained by national libraries, proposing guidelines and strategies based on semantic enrichment to enhance the quality of this representation in information systems. It is a qualitative investigation, with a descriptive and exploratory approach, based on bibliographic review, documentary analysis, and direct observation of public catalogs. Thirteen national libraries were analyzed, selected for their adoption of different metadata standards, in order to identify the fields used to represent gender, the values assigned, the controlled vocabularies employed, as well as the eventual linkage to external identifiers such as Wikidata and VIAF. The study also examined institutional guidelines, technical documents, and public policies related to the inclusion of gender metadata, when available. The results indicate that although metadata standards already provide specific fields for recording gender, their application remains limited, predominantly binary, and lacking standardization across the institutions analyzed. A lack of clear institutional policies for the use of these fields was observed, as well as a predominance of descriptive practices that overlook self-identification and the fluidity of gender identities. On the other hand, positive practices were identified, including the use of diversity-aware controlled vocabularies, linking to external URIs, and records aligned with open ontologies. It is concluded that the qualification of gender metadata in authority records, through the use of linked data technologies and semantic enrichment, can enhance the expressiveness, interoperability, and informational justice of bibliographic catalogs. The adoption of technical and ethical guidelines for such representation is essential to ensure the visibility of women authors, trans and non-binary individuals, and other identities frequently underrepresented in traditional systems. The research thus highlights the importance of normative and political advances to foster diversity and equity in informational environments.

Keywords: Authority records; Gender; Metadata; Semantic Web; Semantic enrichment.

LISTA DE FIGURAS

Figura 1 - Modelo conceitual para dados de autoridade	39
Figura 2 - Pontos de acesso	47
Figura 3 - Documentos recuperados por bases de dados	57
Figura 4 - Documentos recuperados, aceitos e rejeitados	58
Figura 5 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional Mariano Moreno	88
Figura 6 - Registro de autoridade de Virginia Woolf no catálogo da Biblioteca Nacional da Colômbia	90
Figura 7 - Registro de autoridade de Virginia Woolf no catálogo da Biblioteca Nacional da Colômbia	91
Figura 8 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional do México	92
Figura 9 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional do Peru	93
Figura 10 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Nacional do Peru	94
Figura 11 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Nacional Alemanha	95
Figura 12 - Registro MARC 21 XML de autoridade da autora Jane Austen no catálogo da Biblioteca Nacional da Alemanha	97
Figura 13 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Real da Bélgica	98
Figura 14 - Separação por gênero no catálogo da Biblioteca Real da Bélgica	99
Figura 15 - Registro MARC 21 de autoridade da autora Jane Austen no catálogo da Biblioteca Nacional de España	100
Figura 16 - Autores en la BNE	101
Figura 17 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional da República Checa	102
Figura 18 - Registro de autoridade de J. L. Butler no catálogo da Biblioteca Nacional da Islândia	103
Figura 19 - Registro MARC 21 de autoridade de J. L. Butler no catálogo da Biblioteca Nacional da Islândia	104
Figura 20 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Nacional da França	105
Figura 21 - Registro UNIMARC de autoridade de Jane Austen no catálogo da Biblioteca Nacional da França	106
Figura 22 - Registro INTERMARC de autoridade de Jane Austen no catálogo da Biblioteca Nacional da França	107
Figura 23 - Registro CONMARC/A de autoridade de Agatha Christie na documentação técnica da Biblioteca Nacional da Eslovênia	108
Figura 24 - Registro MARC 21 de autoridade de Lola Anglada no Catálogo de autoridades CANTIC	109
Figura 25 - Registro de autoridade na interface pública da autora Jane Austen no Biblioteca Nacional da Noruega	110
Figura 26 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional da Noruega	111

LISTA DE QUADROS

Quadro 1- Dados estruturados, semiestruturados e não estruturados	26
Quadro 2 - Definição dos tipos de metadados	28
Quadro 3 - Campos MARC 21 para dados de autoridade	34
Quadro 4 - Novos campos 3XX	34
Quadro 5 - RDA/FRAD Mapeamento de atributos da entidade Pessoa	43
Quadro 6 - Comparativo entre FRAD, RDA e MARC 21 de autoridade	44
Quadro 7 - Classificação dados abertos	64
Quadro 8 - Artigos que abordam o enriquecimento semântico de metadados em LAM	70
Quadro 9 - Tecnologias Semânticas	77
Quadro 10 - Bibliotecas Nacionais que utilizam ou não o metadado de gênero	81
Quadro 11 - Bibliotecas que utilizam o metadado de gênero e técnicas semânticas	112
Quadro 12 – Diretrizes para a inclusão do metadado de gênero em registros de autoridade	119

LISTA DE ABREVIATURAS E SIGLAS

AACR2	Anglo-American Cataloguing Rules, Código de Catalogação Anglo-Americano
BIBFRAME	Bibliographic Framework
BNC	Biblioteca Nacional de Colombia
BNE	Biblioteca Nacional de España
BNMM	Biblioteca Nacional Mariano Moreno
BNP	Biblioteca Nacional del Perú
CI	Ciência da Informação
DC	Dublin Core Metadata Element Set
DNB	Biblioteca Nacional da Alemanha
EDM	Europeana Data Model
FBN	Fundação Biblioteca Nacional do Brasil
FRAD	Functional Requirements for Authority Data
FRBR	Functional Requirements for Bibliographic Records
FRSAD	Functional Requirements for Subject Authority Data
FRANAR	Functional Requirements and Numbering of Authority Records
GARR	Guidelines for Authority Records and References
GND	Gemeinsame Normdatei
HQs	Histórias em quadrinhos
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IFLA	International Federation of Library Associations and Institutions
IFLA LRM	International Federation of Library Associations and Institutions Library Reference Model
ISAAR(CPF)	International Standard Archival Authority Record for Corporate Bodies, Persons and Families
ISBD	International Standard Bibliographic Description
ISNI	International Standard Name Identifier
KOS	Knowledge Organization Systems
LAM	Bibliotecas, Arquivos e Museus (Libraries, Archives and Museums)

LC	Library of Congress
LCNAF	Library of Congress Name Authority File
LCSH	Library of Congress Subject Headings
LD	Linked Data
LOD	Linked Open Data
MARC 21	Machine-Readable Cataloging 21
MLAR	Mandatory Data Elements for Internationally Shared Resource Authority Records
MODS	Metadata Object Description Schema
NACO	Name Authority Cooperative Program
ODS	Objetivos de Desenvolvimento Sustentável
OPACs	Online Public Access Catalogs
OWL	Ontology Web Language
PCC	Program for Cooperative Cataloging
PIC	Declaração de Princípios Internacionais de Catalogação
RDA	Resource Description and Access
RDF	Resource Description Framework
SPARQL	Simple Protocol and RDF Query Language
URI	Unified Resource Identifier
VIAF	Virtual International Authority File
WWW	World Wide Web
XML	Extensible Markup Language

SUMÁRIO

1 INTRODUÇÃO	13
1.1 Justificativa	18
1.2 Objetivos	19
1.3 Estrutura do trabalho	20
2 REPRESENTAÇÃO, ORGANIZAÇÃO DA INFORMAÇÃO E METADADOS	22
2.1 Conceitos fundamentais sobre dados e metadados	25
2.2 Padrões de metadados	32
2.3 Modelos conceituais e o RDA Oficial e Original	37
2.4 Registro de autoridade	46
2.5 Catalogação e a inclusão de gênero	50
3 PROCEDIMENTOS METODOLÓGICOS	55
3 WEB SEMÂNTICA E LINKED DATA	61
3.1 Enriquecimento de metadados LAM	67
3.2 Enriquecimento semântico de metadados LAM	70
4. RESULTADOS E DISCUSSÃO	80
4.1 Levantamento sobre a utilização do metadado de gênero em registros de autoridade	80
4.2 Iniciativas dos catálogos que incluem o metadado de gênero	87
4.2.1 Biblioteca Nacional Mariano Moreno (BNMM)	88
4.2.2 Biblioteca Nacional da Colômbia	90
4.2.3 Biblioteca Nacional do México	91
4.2.4 Biblioteca Nacional do Peru	93
4.2.5 Biblioteca Nacional da Alemanha	95
4.2.6 Biblioteca Real da Bélgica	97
4.2.7 Biblioteca Nacional da Espanha	99
4.2.8 Biblioteca Nacional da República Checa	102
4.2.9 Biblioteca Nacional da Islândia	103
4.2.10 Biblioteca Nacional da França	104
4.2.11 Biblioteca Nacional da Eslovênia	107
4.2.12 Catálogo de autoridades de nomes e títulos da Catalunha	108
4.2.13 Biblioteca Nacional da Noruega	110
5. BOAS PRÁTICAS A DETERMINAÇÃO DO METADADO DE GÊNERO	117
6 CONSIDERAÇÕES FINAIS	121
REFERÊNCIAS	126

1 INTRODUÇÃO

Os sistemas de informação coletam, tratam, organizam, armazenam e disseminam o conhecimento. A catalogação influencia a representação e recuperação das informações nos catálogos, estruturando dados bibliográficos e de autoridade de maneira sistemática, seguindo regras e padrões. O avanço das tecnologias da informação (TIC) transformou esse ambiente, ampliando as possibilidades de organização e acesso aos dados. Esse cenário trouxe novos desafios às descrições nos catálogos, exigindo adaptações para atender às necessidades dos usuários, garantir interoperabilidade entre sistemas e melhorar a recuperação da informação.

De acordo com a Comissão Europeia, a tecnologia "reflete os valores dos seus desenvolvedores e os da informação a que recorrem" (European Commission, 2018, não paginado, tradução nossa). A Unesco (2019), por meio do relatório "*I'd blush if i could*", ressalta a importância de equipes com mais diversidade e a capacitação de mulheres para se tornarem criadoras digitais. Essa relação se torna evidente quando se considera a representação do gênero nos registros de autoridade, visto que a forma como a informação é descrita pode influenciar sua visibilidade e acesso.

Nos últimos anos, a conscientização sobre a importância da igualdade de gênero tem se ampliado em diversas áreas, incluindo a Ciência da Informação. Essa preocupação está alinhada a iniciativas globais, como os Objetivos de Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas (ONU), especificamente o Objetivo 5, que trata da igualdade de gênero (ONU, 2015). No âmbito da catalogação, garantir a representação adequada do gênero nos registros de autoridade contribui para a recuperação da informação, inclusão e visibilidade.

Arakaki, F., Simionato e Santos (2017, p. 4) destacam que a catalogação "[...] está alinhada às tecnologias de informação, no que diz respeito do uso e reuso de registros em bibliotecas para a construção de representações que possibilitem aos usuários encontrar, identificar, selecionar, adquirir ou obter e navegar no catálogo". Nesse sentido, a catalogação acompanha as transformações tecnológicas ao incorporar novas práticas que ampliam a integração dos registros bibliográficos e de autoridade em diferentes ambientes digitais.

Os catálogos passaram por diversas transformações, desde inventários manuais até sistemas eletrônicos, como os *Online Public Access Catalogs* (OPACs), que mudaram a forma de realizar pesquisas e a recuperação da informação. Essa

evolução exige a constante adaptação dos catálogos para atender às novas demandas informacionais. Como destaca a *International Federation of Library Associations and Institutions* (IFLA, 2016, p. 4), é essencial que os catálogos sejam ajustados às "[...] novas categorias de usuários, o entorno do acesso aberto, a interoperabilidade e acessibilidade aos dados, as características das ferramentas de descobertas e, em geral, as significativas mudanças no comportamento dos usuários". Além disso, assumem "[...] status de ambiente informacional criado para a busca e a estruturação do conhecimento produzido; e a Catalogação Descritiva passa a ser tratada como uma disciplina que vai além da técnica [...]" (Jesus, 2021, p. 35)

Existe uma diversidade de catálogos, cada um estruturado para atender a necessidades específicas, como temáticos, coletivos e de autoridade. Os catálogos de autoridade, em particular, reúnem nomes de entidades, títulos e assuntos. Instituições ao redor do mundo mantêm seus próprios catálogos de autoridade, cada um com diretrizes e escopos próprios, como o *Library of Congress Name Authority File* (LCNAF), o *Gemeinsame Normdatei* (GND), o Catálogo de Autoridades da *Biblioteca Nacional de España* (BNE), a Rede Bibliodata e o Catálogo de Autoridades da Fundação Biblioteca Nacional do Brasil (FBN).

Para que o catálogo de autoridade tenha consistência, é necessário seguir padrões que garantam a uniformidade na identificação das entidades. Códigos ou esquemas de descrição, como o Código de Catalogação Anglo-Americano (AACR2r), estabelece regras específicas para a descrição bibliográfica, enquanto o *Resource Description and Access* (RDA) fornece diretrizes mais flexíveis e alinhadas a modelos conceituais, incluindo atributos como o gênero, de forma opcional.

Além disso, o controle de autoridade nos catálogos normaliza os pontos de acesso, evita ambiguidades causadas por homônimos ou variações de nomes e promove a recuperação da informação, estabelecendo conexões entre diferentes formas de acesso, assegurando coerência e interoperabilidade. Esse processo envolve diversas etapas que resultam na criação de um registro de autoridade, no qual é estabelecida uma forma padronizada do nome a ser adotada em todos os registros do catálogo (Herrero Pascual, 1999, p. 1).

Os registros de catalogação dependem da utilização de metadados, Gilliland (2016, p. 2, tradução nossa) define metadados como "[...] informações de valor agregado que criam para organizar, descrever, rastrear e melhorar o acesso a objetos de informação e itens físicos e coleções, relacionados a esses objetos". Essa

abordagem evidencia que os metadados descrevem, estruturam e facilitam o gerenciamento da informação, tornando-a mais acessível em diferentes contextos.

Metadados, isoladamente, não asseguram consistência e interoperabilidade, devem ser estruturados com base em padrões. Esses padrões fornecem critérios, métodos e processos padronizados que orientam a concepção, desenvolvimento e implementação das estruturas de dados, seus valores, conteúdos e a troca de informações de forma eficiente (Zeng; Qin, 2016, p. 23).

Entre os principais padrões utilizados para a descrição bibliográfica estão o *Machine-Readable Cataloging 21* (MARC 21), o *Dublin Core Metadata Element Set* e o *Bibliographic Framework* (BIBFRAME) (Alves, 2018). No caso do MARC 21 de autoridade, desde suas primeiras especificações já estão previstos elementos voltados à identificação de aspectos pessoais e sociais, como o campo 375, destinado à indicação do gênero da pessoa representada. Apesar de sua existência no padrão, o uso desse campo ainda é pouco frequente nos registros de autoridade, o que revela uma lacuna entre as possibilidades técnicas do formato e as práticas efetivas de catalogação.

Nesse contexto, o *Program for Cooperative Cataloging* (PCC) formou um Grupo de Trabalho *Ad Hoc* sobre Gênero em Registros de Autoridade de Nomes, que emitiu um relatório em 2016 abordando a inclusão de informações de gênero nesses registros. O relatório recomendou a utilização do campo do MARC 21, 375 para registrar o gênero, empregando termos específicos e controlados, como "feminino", "masculino" ou "desconhecido". No entanto, essa prática gerou debates sobre privacidade e a complexidade das identidades de gênero. Em resposta às preocupações levantadas, o PCC revisou suas diretrizes e, em 2022, publicou um relatório atualizado recomendando que o elemento de gênero não seja mais registrado em registros de autoridade de nomes pessoais (PCC Ad Hoc Task Group on Gender in Name Authority Records, 2016; 2022).

Diante disso, Possemato (2025) analisa a posição do PCC e do *International Standard Name Identifier* (ISNI) em retirar o gênero. Destaca que a catalogação está se transformando com a modelagem de entidades, que prioriza a identificação das relações entre entidades. Atributos antes considerados fixos, como o gênero, passam a ser reavaliados de forma contextualizada nos registros de autoridade. A catalogação tradicional tratava o gênero como uma propriedade intrínseca e estática, sem considerar sua fluidez e construções socioculturais.

A *Web Semântica* e o *Linked Data*, possibilitam uma representação mais dinâmica e conectada dos dados. A *Web Semântica*, proposta por Berners-Lee, Hendler e Lassila (2001, tradução nossa), é "[...] uma extensão da atual, onde a informação tem um significado bem definido, permitindo que os computadores e as pessoas trabalhem em cooperação", o que expande as possibilidades de organização e recuperação da informação ao conectar os dados de maneira mais expressiva.

Esses novos formatos para representar os dados levaram Berners-Lee (2006) a criar o conceito de *Linked Data*, definido como um conjunto de práticas para publicar dados estruturados na *Web*, permitindo que informações sejam interligadas e reutilizadas por diferentes sistemas. Essa proposta amplia as possibilidades de interoperabilidade entre domínios e fomenta a construção de redes semânticas abertas e conectadas.

Nesse contexto, a aplicação das tecnologias de *Linked Data* à representação informacional de registros oriundos de Bibliotecas, Arquivos e Museus (LAM) demanda uma abordagem que vá além da simples conversão de formatos: exige o enriquecimento semântico dos metadados. Como destaca Zeng (2019, p. 07, tradução nossa), o enriquecimento de metadados LAM tem "o objetivo de superar desafios relacionados à qualidade dos dados e à sua capacidade de descoberta na era digital, ao mesmo tempo em que fornece mais contexto e informações multilíngues para objetos do patrimônio cultural (CH)".

No estudo feito por Peng *et al.* (2015), observou-se que as informações básicas, como data de publicação, autor e idioma, embora essenciais, não atendiam a todas as demandas informacionais dos pesquisadores. Por exemplo, quando os pesquisadores buscavam estudar estilos de escritas de autores femininos e masculinos no século XIX, enfrentam dificuldades em determinar o gênero dos autores, pois essa informação não se apresenta nos metadados. Para solucionar esse problema utilizaram três fontes diferentes para enriquecer os metadados e por meio dessa abordagem, os pesquisadores alcançaram uma taxa de identificação de gênero de 78,9% dos autores presentes na biblioteca digital *HathiTrust*. Os resultados obtidos demonstram claramente o potencial significativo dessa abordagem para melhorar a representação e a recuperação das informações.

Esta pesquisa não tem como foco a discussão sobre o conceito de gênero, mas sim a inclusão desse metadado nos sistemas de informação. O gênero é compreendido como a forma pela qual o indivíduo se identifica, de acordo com Scott

(1995), é uma construção social que transcende o sexo biológico. Ainda Butler (1990) argumenta que o gênero não constitui uma identidade fixa. Em vez disso, ela sustenta que o gênero é expresso por meio da linguagem e de práticas culturais. Enquanto o sexo se refere às características biológicas que distinguem homens e mulheres, o gênero abrange identidades, comportamentos, expectativas sociais e papéis atribuídos com base nessas diferenças.

Embora o estudo trate da presença do metadado de gênero na catalogação de forma abrangente, os exemplos analisados nos registros de autoridade concentram-se na representação do gênero feminino. Essa delimitação está relacionada ao percurso da pesquisa da própria autora, que iniciou seus estudos sobre a inclusão do gênero em registros de autoras na *Europeana*, no Trabalho de Conclusão de Curso da Graduação em Biblioteconomia na Universidade de Brasília (UnB)¹, com foco na representação de mulheres e identidades femininas no âmbito do patrimônio cultural.

É importante mencionar que apesar da existência de padrões consolidados para a representação de atributos pessoais nos registros de autoridade, como o campo 375 no formato MARC 21 e o campo 120 no UNIMARC, observa-se que o metadado de gênero ainda é pouco utilizado ou aplicado de forma limitada pelas instituições bibliográficas. Embora o tema tenha ganhado certa visibilidade em debates recentes, principalmente em decorrência das discussões promovidas por grupos como o *Program for Cooperative Cataloging* (PCC), a literatura especializada na área de Ciência da Informação ainda carece de investigações aprofundadas que articulem a presença do metadado de gênero às práticas efetivas de catalogação e às implicações sociotécnicas envolvidas.

Como visto nos estudos anteriores, as abordagens predominantes concentram-se nos aspectos técnicos dos formatos de metadados e nas diretrizes descritivas, mas pouco exploram as consequências da ausência dessa informação para a visibilidade, a recuperação e a representação de identidades de gênero diversas nos sistemas informacionais. Além disso, são escassos os estudos empíricos que examinam como as bibliotecas nacionais têm implementado, ou não, o registro dessa informação, quais valores utilizam e quais critérios orientam suas decisões institucionais.

¹ O Trabalho de Conclusão de Curso intitulado “O linked data e a Europeana: uma análise da estrutura de dados de uma coleção de obras de arte de artistas femininas” de autoria de Ana Carolina Novaes de Mendonça e orientado por Dr. Felipe Augusto Arakaki está disponível em: <https://bdm.unb.br/handle/10483/33154>.

Essa lacuna revela a necessidade de compreender não apenas os dispositivos técnicos disponíveis para o registro do gênero, mas também os fatores institucionais, éticos e tecnológicos que influenciam sua adoção. Nesse sentido, a presente pesquisa delimita-se ao exame da presença e do tratamento do metadado de gênero nos registros de autoridade de bibliotecas nacionais, buscando identificar as diretrizes existentes, os padrões aplicados e o potencial de uso de tecnologias de enriquecimento semântico como estratégia para qualificar a representação da identidade de gênero de forma mais inclusiva, interoperável e contextualizada.

Parte-se da hipótese de que a inclusão qualificada do metadado de gênero nos registros de autoridade pode ampliar a visibilidade de identidades de gênero diversas, melhorar os processos de recuperação da informação e fortalecer a representatividade nos catálogos bibliográficos. Considera-se ainda que, para além da presença formal desse metadado, sua eficácia depende da adoção de diretrizes que articulem padrões descritivos, vocabulários controlados e políticas institucionais comprometidas com a autodeclaração e a diversidade identitária. Ademais, pressupõe-se que o uso de tecnologias baseadas em *Linked Data* e enriquecimento semântico contribui para tornar esses registros mais granulares, interoperáveis e sensíveis aos contextos socioculturais, permitindo sua vinculação a fontes externas confiáveis, como *Wikidata*, VIAF e ontologias específicas de gênero.

1.1 Justificativa

A inclusão do gênero nos registros de autoridade está inserida em um contexto de transformação das abordagens da catalogação, impulsionado pela adoção das bibliotecas às tecnologias da *Web* e à modelagem de entidades. A maneira como as bibliotecas incorporam ou não essa informação nos registros de autoridade pode impactar na recuperação e visibilidade nos sistemas informacionais.

Esta pesquisa se justifica pela contribuição ao aprimoramento das práticas de catalogação e da representação de gênero nos registros de autoridade. Inserida na área da Ciência da Informação e na Biblioteconomia, está alinhada à Linha de Pesquisa 2 – Tecnologia, Informação e Representação do PPGCI/UFSCar, ao analisar a catalogação e sua relação com metadados, tecnologia e inclusão de gênero nos registros de autoridade.

Sob uma perspectiva social, a proposta está alinhada aos Objetivos de Desenvolvimento Sustentável (ODS), particularmente o objetivo número 5 que diz respeito à igualdade de gênero. A desigualdade de gênero no ambiente informacional reflete um problema mais amplo de invisibilidade de determinados grupos na produção e disseminação do conhecimento.

No âmbito prático, a pesquisa contribui para o aprimoramento das práticas de catalogação e recuperação da informação. A ausência da informação de gênero nos registros de autoridade em catálogos limita a realização de estudos de gênero, como evidenciado por Peng *et al.* (2015). A adoção de técnicas semânticas para o enriquecimento de metadados pode reduzir o retrabalho dos catalogadores, apoiar a tomada de decisão e aprimorar a navegação nos catálogos.

A motivação pessoal da pesquisa está relacionada à trajetória acadêmica da autora, que em estudos anteriores, analisou os metadados da coleção de obras de arte de artistas femininas na *Europeana*. A experiência adquirida revelou lacunas e impulsionou a investigação sobre a presença do metadado de gênero nos registros de autoridade das bibliotecas nacionais. Dessa forma, o estudo amplia o escopo da análise da representação do gênero feminino, agora voltando-se para um ambiente bibliográfico.

Portanto, esta pesquisa propõe uma abordagem que combina perspectivas teóricas, sociais e práticas para enfrentar as lacunas de representatividade de gênero nos sistemas de informação. Ao investigar a presença e diretrizes das bibliotecas nacionais para na inclusão do gênero nos registros de autoridade, a pesquisa contribui para a reflexão sobre os impactos da catalogação e processos para tornar os registros mais granulares utilizando o enriquecimento semântico.

1.2 Objetivos

O objetivo geral desta pesquisa é analisar a presença e o tratamento do metadado de gênero nos registros de autoridade de bibliotecas nacionais, com vistas à proposição de diretrizes e ao uso de recursos de enriquecimento semântico que favoreçam uma representação mais inclusiva e contextualizada da identidade de gênero.

Como objetivos específicos:

- Analisar os fundamentos dos metadados e os padrões de metadados aplicados aos registros de autoridade
- Examinar os fundamentos conceituais e normativos dos metadados aplicados aos registros de autoridade, com foco nos elementos descritivos e estruturais relacionados à identidade de gênero;
- Mapear a ocorrência e as formas de registro do metadado de gênero nos catálogos de autoridade de bibliotecas nacionais;
- Identificar diretrizes, políticas institucionais e práticas adotadas para a inclusão do metadado de gênero, considerando aspectos técnicos, éticos e sociais;
- Propor diretrizes e estratégias de enriquecimento semântico, baseadas em tecnologias de dados conectados (*Linked Data*), para a representação qualificada da identidade de gênero nos registros de autoridade.

1.3 Estrutura do trabalho

Esta pesquisa está estruturada em seis seções principais. A **Seção 1 – Introdução** apresenta a contextualização do tema, delineando a relevância do estudo sobre o metadado de gênero nos registros de autoridade no campo da Biblioteconomia e da Ciência da Informação. São discutidas as motivações que justificam a investigação, os objetivos gerais e específicos, bem como os procedimentos metodológicos adotados.

Na **Seção 2 – Representação da informação e metadados**, desenvolve-se o referencial teórico sobre os fundamentos da representação da informação e dos metadados, abordando conceitos essenciais, padrões descritivos, tipologias, modelos conceituais e formatos utilizados nos registros de autoridade. A seção dá ênfase ao formato MARC 21, à prática da catalogação e à presença do elemento “gênero”, discutindo como a identidade de gênero tem sido tratada nos códigos de catalogação, nas diretrizes institucionais e nos próprios registros de autoridade. São também analisados os debates contemporâneos, avanços alcançados e os desafios ainda presentes nesse campo.

A **Seção 3 – Web Semântica e *Linked Data*** trata das bases conceituais e tecnológicas da Web Semântica e do *Linked Data*, destacando os padrões envolvidos e suas aplicações nos domínios das Bibliotecas, Arquivos e Museus (*Libraries*,

Archives, and Museums - LAM). Essa seção aprofunda o debate sobre o potencial dessas tecnologias para o enriquecimento semântico de metadados, ressaltando suas contribuições para a qualificação descritiva, a interoperabilidade entre sistemas e a representação mais contextualizada de atributos como a identidade de gênero.

Na **Seção 4 – Análise e Resultados**, são apresentados os achados empíricos da pesquisa, com a descrição das práticas adotadas por bibliotecas nacionais em relação ao metadado de gênero nos registros de autoridade. A análise inclui a identificação dos catálogos que registram esse metadado, a forma como os campos são estruturados, os padrões utilizados e o nível de padronização e clareza na representação da informação.

A **Seção 5 – Proposta de boas práticas** apresenta diretrizes para a representação da identidade de gênero em registros de autoridade, com base nos dados analisados e no estado da arte da literatura. A proposta articula aspectos técnicos, sociais e tecnológicos, visando contribuir para uma representação mais inclusiva, ética e interoperável nos sistemas de informação.

Por fim, a **Seção 6 – Considerações finais** retoma os principais achados da pesquisa, destacando suas contribuições teóricas e práticas para a área da Ciência da Informação e da Biblioteconomia. A seção final também aponta limitações do estudo e sugere caminhos para futuras investigações sobre a interseção entre metadados, identidade de gênero e representação da informação em ambientes digitais e semânticos.

2 REPRESENTAÇÃO, ORGANIZAÇÃO DA INFORMAÇÃO E METADADOS

A representação da informação é um dos eixos da Ciência da Informação (CI), que envolve a estruturação e descrição para facilitar sua recuperação e uso. Desde suas origens, a CI tem se caracterizado como um campo interdisciplinar, com relações com as áreas da Ciência da Computação, Comunicação, Linguística e Biblioteconomia. Seu desenvolvimento está ligado à necessidade de tornar a informação acessível e útil para diferentes comunidades de usuários (Borko, 1968).

A necessidade de representar a informação tornou-se mais evidente com o crescimento exponencial da produção de documentos e dados em diversos formatos. Entende-se a informação a partir do conceito de Buckland (1991), onde o autor define a informação como processo que tem como finalidade comunicar algo, como conhecimento a fim de reduzir incertezas e gerar conhecimento e por último como coisa, podendo ser atribuída a qualquer dado, objeto ou documento.

Marcondes (2001, p. 61) aponta que "De nada adianta a informação existir, se quem dela necessita não sabe da sua existência, ou se ela não puder ser encontrada". No contexto digital e marcado pelo excesso informacional, a dificuldade não está apenas na produção de novos conhecimentos, mas na sua visibilidade e acesso. Isso reforça a importância dos sistemas de recuperação da informação e metadados bem estruturados que possibilitem que os usuários encontrem a informação desejada.

A falta de uma informação pode ser resultado de barreiras linguísticas, tecnológicas ou socioeconômicas. Assim, garantir que a informação seja acessível a diferentes públicos implica pensar em políticas de acesso aberto, interoperabilidade entre sistemas e estratégias de divulgação que ampliem a visibilidade dos conteúdos informacionais. A inclusão da informação sobre gênero nos registros de autoridade pode contribuir para a visibilidade da produção intelectual e a organização de coleções. Por outro lado, sua ausência pode dificultar a recuperação dessas obras e impactar análises sobre representatividade na ciência e na cultura.

Nesse sentido, a evolução da CI trouxe novos paradigmas para a representação da informação, ampliando seu escopo para além dos ambientes tradicionais de catalogação. Com o advento da *Web Semântica* e das tecnologias relacionadas, a organização da informação passou a incorporar modelos conceituais, permitindo maior interoperabilidade e enriquecimento semântico. Nesse cenário, os metadados são fundamentais para a descrição e interconexão de recursos

informacionais, possibilitando a integração entre diferentes sistemas e ampliando a acessibilidade da informação.

A catalogação consiste no processo de organizar, descrever e registrar recursos informacionais de forma padronizada e sistemática para que possam ser recuperados posteriormente. Garrido Arilla (1999, local. 522, tradução nossa), pontua que catalogar é “[...] descrever os elementos informativos que permitem identificar um documento e estabelecer os pontos de acesso que possibilitem recuperá-lo pelo título, autor ou assunto previamente conhecidos.”

Além de fornecer acesso à informação, a catalogação estrutura os registros informacionais, contribuindo para a gestão do conhecimento em ambientes bibliográficos, arquivísticos e museológicos (LAM). Santos (2013, não paginado) discorre que a catalogação

Responsável pela construção de sistemas constituídos por conjuntos de elementos interconectados e combinados de formas de representação cria instrumentos para facilitar o fluxo da informação nos mais diversos ambientes informacionais é responsável pela apresentação de estruturas que oferecem condições favoráveis de acesso aos códigos formais de representação simbólica e aos canais de transferência da informação, desempenhando com competência a decodificação e a codificação dos códigos e regras utilizados para representar conhecimento e descrever informações, documentos e recursos.

O propósito da catalogação segundo Guinchat e Menou (1994, p. 101) é “[...] fornecer uma representação do documento que é descrito de uma forma única e não ambígua, o que permite identificá-lo, localizá-lo, representá-lo nos catálogos correspondentes e recuperá-lo.”

Existem 13 princípios gerais de catalogação estabelecidos na Declaração de Princípios Internacionais de Catalogação, que visam promover uma abordagem consistente e centrada na conveniência do usuário. Os princípios são, interesse do usuário, uso comum, representação; precisão, suficiência e necessidade; significantes, economia, coerência e a normalização, integração, interoperabilidade, abertura, acessibilidade e racionalidade (IFLA, 2016).

Entre os aspectos os mais relevantes para essa pesquisa, destacam-se os princípios de representação, precisão, significação, normalização, interoperabilidade e acessibilidade, que desempenham um papel significativo na inclusão do gênero nos registros de autoridade.

A representação sugere que as entidades sejam descritas conforme sua forma mais reconhecida ou autodesignada, garantindo que os registros reflitam com

precisão a identidade de pessoas, entidades coletivas e obras catalogadas. No contexto da catalogação do gênero, esse princípio pode contribuir para uma representação que evita distorções ou omissões nos registros de autoridade.

Já os princípios de coerência e normalização, estão ligados a normalização das descrições e criação de pontos de acesso. No caso do gênero, a normalização pode ser aplicada por meio de regras que orientem a inclusão desse atributo nos registros de autoridade, utilizando terminologias padronizadas que promovam a interoperabilidade entre sistemas bibliográficos.

A significação os elementos devem conter informações pertinentes à descrição, suficientemente relevantes para serem incluídas e capazes de distinguir uma entidade de outra. A interoperabilidade possibilita a conexão e reutilização dos registros em diferentes bases, garantindo que o metadado de gênero, quando incluído de maneira estruturada, favoreça pesquisas mais abrangentes sobre a produção intelectual e formação de coleções. Já a acessibilidade visa assegurar que os registros sejam compreensíveis e recuperáveis por diferentes públicos, promovendo maior diversidade nos sistemas informacionais.

A aplicação desses princípios pode contribuir para evitar a invisibilização do gênero nos registros de autoridade, garantindo que as informações sejam descritas de maneira fiel e equitativa. Além disso, registros que respeitam a autodescrição da entidade ampliam a diversidade nos sistemas de informação e promovem um compartilhamento mais representativo dos dados, melhorando a recuperação da informação.

O processo de catalogação pode ser dividido em tipos específicos, que se complementam e atendem às diferentes necessidades de gestão de recursos informacionais. Dessa forma, Read (2003) aponta que a catalogação pode ser classificada como descritiva ou de assunto.

Para Garrido Arilla (1996) a catalogação descritiva é um processo que começa com a descrição bibliográfica do documento, segue com a seleção de pontos de acesso e conclui com a organização e registro das informações locais atribuídas ao material. Silva e Fujita (2004) discorrem que a catalogação de assunto é a atribuição de cabeçalhos que representam o conteúdo nos catálogos de bibliotecas e são construídos como enunciados organizados a partir da composição ordenada de palavras.

Em ambas as abordagens o julgamento do catalogador influencia nas decisões na criação dos registros do catálogo, o que impacta o acesso da informação. As escolhas podem destacar ou ocultar materiais, conferindo ao catalogador uma influência significativa sobre o acesso ao conhecimento (Martin, 2021). A seguir, serão discutidos os diferentes tipos de dados e os metadados.

2.1 Conceitos fundamentais sobre dados e metadados

Os dados são unidades elementares que representam fatos ou observações de forma simbólica. De acordo com Santos e Sant'Ana (2013, p. 205) dado é

[...] uma unidade de conteúdo necessariamente relacionada a determinado contexto e composta pela tríade entidade, atributo e valor, de tal forma que, mesmo que não esteja explícito o detalhamento sobre contexto do conteúdo, ele deverá estar disponível de modo implícito no utilizador, permitindo, portanto, sua plena interpretação.

Santos e Sant'Ana (2013) enfatizam sua natureza contextual e relacional, destacando que ele não existe isoladamente, mas sempre vinculado a uma entidade específica, caracterizada por atributos e valores. Essa estrutura permite que os dados sejam interpretáveis, mesmo quando o contexto completo não está explicitamente mencionado, pois o utilizador pode inferi-lo a partir de seu conhecimento prévio.

Arakaki, A. e Arakaki, F. (2020) discutem as relações conceituais entre os termos dado e metadado no âmbito da Ciência da Informação, destacando que a interdisciplinaridade com áreas como Computação, Ciência de Dados e Web Semântica tem gerado confusões terminológicas. Os autores argumentam que não há distinções fixas entre dado e metadado, pois essas categorias dependem do contexto de uso e da perspectiva adotada. O mesmo registro pode ser considerado metadado, quando utilizado para descoberta de informação, ou dado, quando serve como base analítica ou documental. A análise é embasada no perspectivismo, o que permite observar como diferentes comunidades de prática moldam essas definições. Além disso, os autores reforçam o papel central dos metadados para a curadoria digital, interoperabilidade e processos de Big Data, ressaltando sua importância como insumo fundamental na Ciência de Dados.

A distinção entre diferentes tipos de dados é importante para a representação, organização e recuperação da informação. Schöch (2013) enfatiza essa

diferenciação, destacando que os dados estruturados, frequentemente são armazenados em bancos de dados, onde pares chave/valor possuem identificadores claros e relações bem definidas, seguindo um modelo de dados explícito. Em contraste, dados não estruturados, como texto simples, não possuem limites claros entre os itens, e suas relações e significados geralmente são implícitos.

Por sua vez, os dados semiestruturados, como aqueles armazenados em arquivos *Extensible Markup Language* (XML), situam-se entre os dois extremos, podendo ser mais ou menos rigorosamente organizados dependendo da presença ou ausência de um esquema que os delimite. No Quadro 1 abaixo, estão alguns exemplos de dados estruturados, semiestruturados e não estruturados.

Quadro 1- Dados estruturados, semiestruturados e não estruturados

Tipo de dado	Exemplos de dados
Estruturado	<ul style="list-style-type: none"> ● Catálogos de todos os tipos ● Guias de arquivos ● Bibliografias nacionais ● Bases de dados de indexação e resumos ● Portais de coleções especiais ● Conjuntos de dados de pesquisa curados ● Sistemas de Organização do Conhecimento (KOS)
Semiestruturado	<p>Seções não estruturadas dentro de descrições de metadados</p> <ul style="list-style-type: none"> ● Documentação arquivística não contida no EAD (Encoded Archival Description) ● Arquivos TEI (Text Encoding Initiative), corpo principal fora do cabeçalho ● Recursos com valor agregado ou marcados que existem em todos os tipos de formatos ● Partes não estruturadas de conjuntos de dados estruturados
Não estruturado	<p>Documentos, artefatos culturais e objetos portadores de informação original, que podem ser:</p> <ul style="list-style-type: none"> ● Digitalizados ou não digitalizados ● Textuais ou não textuais ● Em todos os tipos de formatos e mídias ● Possivelmente com data e origem indeterminadas

Fonte: Adaptado e traduzido de Zeng (2019, p. 03).

Como apresentado no Quadro 1, os dados estruturados são utilizados em sistemas de indexação, catálogos de bibliotecas, bases de dados acadêmicas e

sistemas de organização do conhecimento (*Knowledge Organization Systems – KOS*). Entre os diferentes tipos de dados estruturados, aqueles aplicados em registros bibliográficos e de autoridade se destacam por garantir a padronização e a precisão na recuperação da informação. Para que esses dados sejam compreendidos, reutilizados e localizados de maneira eficiente, é necessário que estejam acompanhados de metadados.

O termo "*metadata*" foi introduzido por Jack E. Myers em 1969 e, desde então, tem sido amplamente utilizado em diversas áreas, como Ciência da Computação, Estatística, Bancos de Dados, Biblioteconomia e Ciência da Informação, para se referir a atributos que descrevem, contextualizam e registram características de objetos ou conjuntos de dados (Greenberg, 2005).

De forma simples, metadados são descritos como “dados sobre dados”, porém essa definição não capta a complexidade de seu papel na estruturação, recuperação e uso da informação. Pomerantz (2015, p. 26, tradução nossa) questiona a utilidade dessa definição, argumentando que metadados são, na verdade, “[...] uma declaração sobre um objeto potencialmente informativo”. Essa perspectiva amplia a compreensão dos metadados e destaca seu papel na criação de relações entre objetos e no fornecimento de contexto para sua interpretação.

Já Zeng e Qin (2008, p. 07, tradução nossa) definem que “[...] metadados encapsulam informações que descrevem qualquer documento ou objeto nos formatos digital e tradicional”. Essa definição destaca a abrangência dos metadados, que não se restringem a ambientes digitais, mas também na representação de documentos físicos. Alves (2010, p. 47) define metadados, como

[...] são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação.

A definição de Alves (2010) aponta o papel dos metadados na estruturação da informação, o que melhora a busca e a navegação nos sistemas. Ao serem considerados atributos que representam uma entidade dentro de um sistema, os metadados descrevem um objeto informacional e garantem sua identificação e recuperação de forma única.

Riley (2017) ressalta que os metadados são elementos estruturantes nos sistemas de armazenamento de informação, permitindo que os usuários localizem conteúdos relevantes, registrem dados significativos sobre esses recursos e compartilhem essas informações com outras pessoas.

Os metadados possuem diferentes tipologias e desempenham funções variadas na organização da informação. Essas classificações auxiliam na compreensão de sua estrutura e aplicação em diferentes contextos, permitindo que sejam utilizados de maneira mais eficiente nos sistemas de informação. Diferentes pesquisadores já abordaram a categorização dos metadados, propondo modelos que destacam suas funções e características.

Arakaki, F. (2019) reuniu diversas contribuições teóricas e propôs um conjunto de definições e hierarquia para os metadados, fundamentado em autores como Gilliland (1999, 2008, 2016), Greenberg (2001), Méndez Rodríguez (2002), Senso e Rosa Piñero (2003), Riley (2004), Haynes (2004), Hillman, Marker e Brady (2008), Alves (2010), Miller (2011), Alves e Santos (2013), Pomerantz (2015), Zeng e Qin (2008; 2016), Gartner (2016), Riley (2017), Haynes (2017) e Joudrey e Taylor (2018). Essas abordagens consolidam o entendimento sobre as categorias de metadados e suas aplicações na descrição e recuperação da informação. No Quadro 2, estão apresentados os principais tipos de metadados e suas características.

Quadro 2 - Definição dos tipos de metadados

Tipo	Definição
Metadados Administrativos	Metadados administrativos são usados para gerenciar e administrar coleções e recursos informacionais, para auxiliar na tomada de decisão e manutenção dos registros.
Metadados de autenticação	Metadados de autenticação são informações que possibilitam a identificação, integridade, legitimidade de um recurso informacional. Exemplos consistem em: código de identificação ou verificação, assinatura digital, entre outros. (Greenberg, 2001).
Metadados Preservação	Metadados de preservação estão relacionados com informações de preservação e conservação dos recursos informacionais.
Metadados de Proveniência	Metadados de proveniência estão relacionadas às informações de procedência, fornece dados sobre entidades, criação e modificações e seus relacionamentos. (Pomerantz, 2015).
Metadados Técnicos	Metadados técnicos estão relacionados a como um sistema funciona, fornecendo informações do sistema ou do recurso
Meta-metadada	Meta-metadada corresponde à informações sobre o registro criado, ou informações da criação de um conjunto de dados.

Metadados Descritivos	Metadados descritivos descrevem características identificadoras e os contextos intelectuais dos recursos de informação para fins de descoberta, identificação, seleção, aquisição, contexto e compreensão. (Joudrey; Taylor, 2018).
Metadados de Direitos	Metadados de direitos estão relacionados às informações sobre propriedade, e direitos autorais.
Metadados de acesso e uso	Metadados acesso e uso são informações de como um recurso informacional foi acessado e usado, como restrições de circulação e acesso, registros de exposições, entre outros
Metadados estruturais	Metadados estruturais estão relacionados à composição e organização do recurso informacional.
Markup languages	Markup languages integra metadados e sinalizações para outros recursos estruturais ou semânticos. (Riley, 2017).

Fonte: Arakaki, F. (2019, p. 81).

Os metadados administrativos, como destacado por Arakaki, F. (2019) no Quadro 2, auxiliam na tomada de decisão dentro da gestão de coleções e preservação da informação. A forma como os metadados são estruturados e utilizados pode influenciar diretamente a eficiência dos processos administrativos e técnicos em bibliotecas, arquivos e museus.

A tomada de decisão a partir dos metadados está relacionada à maneira como a informação é organizada, estruturada e disponibilizada nos sistemas de informação. Fornecer dados completos e atualizados permite que as instituições tomem decisões informadas sobre gestão de acervos, preservação digital e acesso à informação. No entanto, existem situações em que as decisões administrativas são frequentemente tomadas sob pressão, com base em informações incompletas, desatualizadas ou fundamentadas na experiência individual (Arakaki, A.; Isotani, 2024; Pfeffer; Sutton, 2006).

Os metadados administrativos podem ajudar a minimizar esses desafios, fornecendo um conjunto estruturado de informações que apoiam a governança documental e garantem a rastreabilidade das decisões institucionais. Segundo Pomerantz (2015, p. 17-18, tradução nossa), os metadados administrativos "[...] fornecem informações sobre a origem e a manutenção de um objeto [...]", o que permite um controle mais preciso dos registros informacionais ao longo do tempo.

Já os metadados de autenticação avaliam a integridade, legitimidade e qualidade dos objetos informacionais. Segundo Greenberg (2001, p. 919, tradução nossa), esses metadados "[...] suportam a avaliação da integridade de um objeto de

informação, legitimidade e qualidade geral genuína". Isso significa que esses metadados fornecem informações que ajudam a verificar se um recurso foi alterado, corrompido ou se mantém sua autenticidade ao longo do tempo.

Riley (2017) discorre que metadados de preservação auxiliam na interoperabilidade e gestão de objetos digitais, garantindo sua longevidade. Esses metadados incluem *checksums*, que verificam a integridade dos arquivos, e eventos de preservação, que registram as ações realizadas para a conservação dos recursos.

Os metadados de proveniência registram a história e o contexto de um recurso informacional, incluindo informações sobre sua criação, modificações e a relação com outras entidades. De acordo com Pomerantz (2015, p. 101, tradução nossa), a "[...] proveniência significa não apenas a história de um recurso, mas as relações entre esse recurso e outras entidades que influenciaram sua história."

Os metadados técnicos conforme o Alves e Santos (2013) são

[...] relacionados com o funcionamento dos sistemas e o comportamento dos metadados. Esse tipo fornece informações sobre hardware e software, digitalização, controle do tempo de resposta dos sistemas, autenticidade e segurança dos dados (criptografia e senhas) etc.

Já os *meta-metadata* "[...] são importantes para garantir a autenticidade dos metadados e acompanhar o processo interno" (Joudrey; Taylor; Wisser, 2018, p. 196, tradução nossa). De forma geral, são metadados que descrevem outros metadados, fornecendo informações sobre o uso, origem, estrutura e gestão dos mesmos.

Riley (2017) explica que os metadados descritivos são amplamente utilizados em instituições de patrimônio cultural. Esses metadados possibilitam a descrição detalhada de recursos informacionais em catálogos, promovem a padronização dos registros e facilitam o acesso aos conteúdos disponíveis. Além disso, contribuem para a identificação e organização dos recursos, pois incluem informações como título, autor, assunto, gênero e data de publicação, auxiliando na busca e na recuperação eficiente da informação.

Segundo Zeng e Qin (2016), os metadados descritivos são utilizados para "[...] identificar e descrever coleções e recursos de informações relacionadas (exemplos incluem registros de catalogação, ferramentas de localização, índices especializados e informações curatoriais." A estruturação permite que usuários localizem e compreendam melhor os objetos descritos, sejam eles livros, documentos

arquivísticos, obras de arte, registros de autoridade ou itens de coleções museológicas.

Os metadados de direitos são informações relacionadas à propriedade intelectual, direitos autorais, restrições de uso e permissões de reprodução de um recurso (Miller, 2011, p. 12). Por exemplo, o *Creative Commons* (CC) e direitos autorais nos registros bibliográficos no campo 540 do MARC21. Além disso, os metadados de acesso e uso podem fornecer informações sobre determinado objeto foi utilizado, podendo sobrepor os metadados de direitos (Miller, 2011; Pomerantz, 2015).

Joudrey, Taylor e Wisser (2018) apontam que os metadados estruturais, incluem informações técnicas para que um recurso digital possa ser acessado e exibido para os usuários. Por fim, o *markup language* são as linguagens de marcação, como o XML.

A classificação e o uso dos metadados vão além de aspectos técnicos, abrangendo também questões éticas, especialmente no que diz respeito à representação da informação e ao impacto dessas descrições na recuperação e disseminação do conhecimento. A ética nos metadados está relacionada aos esquemas de descrição e valores que guiam a forma como as informações bibliográficas são registradas e estruturadas, influenciando diretamente a maneira como os usuários acessam e interagem com os recursos informacionais disponíveis (Sardo, 2019). Fica claro que a questão ética nos metadados não está isolada de escolhas e valores. Pelo contrário, ela emerge justamente nas decisões sobre como os dados são representados, impactando a precisão, a correção e, conseqüentemente, o sucesso das pesquisas.

Quando metadados são gerados com imprecisões ou erros, mesmo que triviais, isso pode distorcer o acesso à informação, tornando o trabalho de catalogação inútil e, em última instância, prejudicando os usuários (Sardo, 2019). Portanto, a ética nos metadados busca garantir que o processo de criação e manutenção dos dados seja feito de forma correta e justa, o que garante o acesso adequado e equitativo à informação.

Para diminuir erros e imprecisões os padrões de metadados são fundamentais, pois assim os sistemas podem compartilhar e organizar de forma padronizada seus metadados. A próxima subseção discorre sobre os padrões de metadados.

2.2 Padrões de metadados

A definição de padrões e vocabulários, juntamente com valores éticos, garante que os recursos informacionais sejam descritos de forma consistente, permitindo que diferentes sistemas compartilhem e interpretem os dados de maneira uniforme. Segundo Alves (2010, p. 47-48) padrões de metadados são

[...] estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados. O objetivo do padrão de metadados é descrever uma entidade gerando uma representação unívoca e padronizada que possa ser utilizada para recuperação da mesma.

Padrões bem estabelecidos contribuem para a organização da informação e permitem a integração entre sistemas distintos. Conforme Zeng e Qin (2016, p. 23, tradução nossa) os padrões de metadados estabelecem “[...] critérios, métodos, processos e práticas uniformes para orientar o projeto, criação e implementação de estruturas de dados, valores de dados, conteúdo de dados e troca de dados de uma maneira eficiente [...]”.

Santos, Simionato e Arakaki, F. (2014, p. 149) discorrem que “No momento da escolha do padrão de metadados, os catalogadores não devem considerar apenas as necessidades da instituição, mas principalmente, as necessidades informacionais dos usuários”. Ou seja, pode-se garantir que os sistemas informacionais sejam funcionais e atendam às demandas informacionais dos usuários.

Riley (2017) destaca que a padronização dos metadados contribui para garantir consistência e interoperabilidade entre sistemas. Dessa forma, a padronização se torna um fator determinante para a eficiência da recuperação da informação, garantindo que os usuários tenham acesso aos recursos informacionais de maneira precisa e confiável.

No domínio bibliográfico existem diversos padrões de metadados, de acordo com Alves (2018, p. 244), os mais utilizados são

[...] padrão ou formato MARC 21 - *Machine-Readable Catalog* (Catalogação Legível por Máquina); MODS – *Metadata Object Description Schema* (Esquema de metadados para a descrição de objeto) e BIBFRAME – *Bibliographic Framework Initiative* (Iniciativa de Framework Bibliográfico).

Dentre esses padrões, o MARC ainda se destaca como um dos formatos mais amplamente adotados para a representação e intercâmbio de registros bibliográficos.

Criado nos anos 1960, possibilitou a automatização da catalogação, permitindo que bibliotecas compartilhassem registros bibliográficos de maneira padronizada (Riley, 2017).

O formato possui diferentes tipos, cada um voltado para uma função específica na catalogação e intercâmbio de dados. Entre eles estão: MARC 21 Bibliográfico, MARC 21 de Autoridade, MARC 21 de Itens, MARC 21 de Classificação e MARC 21 para Informação de Comunidade. Esses formatos padronizam a representação e o compartilhamento de informações sobre recursos bibliográficos, registros de autoridade, exemplares de acervos, sistemas de classificação e dados comunitários, garantindo a acessibilidade e interoperabilidade entre sistemas automatizados (Library of Congress, 2006).

Um registro MARC 21 é composto por três elementos principais, a estrutura do registro, que é padronizada sob a ANSI/NISO Z39.2, adotada nos Estados Unidos, e a ISO 2709, padrão internacional. A designação de conteúdo é definida por cada formato MARC, e o conteúdo dos dados do registro, como o *International Standard Bibliographic Description* (ISBD), *Anglo-American Cataloguing Rules* (AACR) e *Library of Congress Subject Headings* (LCSH) (Library of Congress, 2009). Dentre estas orientações para a catalogação, pode-se citar também o RDA em sua versão *Original*, que foi elaborado em 2010 para substituir o AACR2. Nesta pesquisa, o foco está no formato para dados de autoridade, sendo assim, será aprofundado apenas o MARC 21 de autoridade.

De acordo com Riley (2017, p. 28, tradução nossa)

[...] o formato MARC 21 de Autoridade é utilizado para documentar termos controlados referentes a pessoas, entidades corporativas, títulos de obras, assuntos e gêneros. Esses termos controlados são inseridos nos campos apropriados dos registros bibliográficos MARC 21, garantindo consistência nos registros e facilitando a recuperação da informação.

Dessa forma, o seu uso possibilita a normalização da catalogação, favorecendo a interoperabilidade entre diferentes sistemas e bases de dados. O MARC 21 de Autoridade possui três componentes principais: líder, diretório e campos variáveis, que apresentam semelhanças com os utilizados no MARC 21 Bibliográfico, incluindo a estrutura baseada em indicadores e subcampos. Os campos variáveis são organizados em blocos numéricos, nos quais o primeiro dígito indica a função dos dados contidos no campo, enquanto os demais números identificam o tipo específico

de informação registrada (Salgado *et al.*, 2022). Os campos do MARC 21 de autoridade estão no Quadro 3.

Quadro 3 - Campos MARC 21 para dados de autoridade

Campos MARC 21 autoridade
0XX Números padronizados, números de classificação, códigos
1XX Pontos de acesso (autorizado e referência)
2XX Remissivas complexas “Ver”
3XX Remissivas complexas “Ver também”
4XX Remissivas “Ver”
5XX Remissivas “Ver também”
6XX Notas
7XX Ponto de acesso de ligação
8XX Gráficos alternativos
9XX Implementações locais

Fonte: Salgado *et al.* (2022, p. 09).

Os pontos de acesso autorizados estão no campo 1XX e não podem repetir, já os pontos de acesso não autorizados estão no campo 4XX e podem repetir. Nos blocos 1XX, 4XX, 6XX, 7XX e 8XX, é mantida uma consistência na designação de conteúdo (Library of Congress, 2018).

Seikel e Steele (2011) apontam mudanças que ocorreram no MARC 21 desde a criação e os novos campos introduzidos após o RDA Original. Os campos 3XX foram incorporados para aprimorar a descrição das entidades, facilitando a identificação de pessoas, entidades coletivas e obras. No RDA Original, há diretrizes específicas para o preenchimento desses campos, complementando as regras já estabelecidas no MARC 21 (Assumpção, 2020, p. 135). Os campos estão no Quadro 4, abaixo.

Quadro 4 - Novos campos 3XX

Campo
336 – Tipo de conteúdo (R) ²

² O código de repetibilidade/não repetibilidade (R ou NR):
<https://www.loc.gov/marc/authority/adintr.html>.

368 – Outros atributos de pessoa e entidade coletiva (R)
370 – Local associado (R)
371 – Endereço (R)
372 – Campo de atividade (R)
373 – Grupo associado (R)
374 – Ocupação (R)
375 – Gênero (R)
376 – Informação da família (R)
377 – Idioma associado (R)
378 – Forma completa do nome pessoal (NR)
380 – Forma da obra (R)
381 – Outra característica distintiva da obra ou expressão (R)

Fonte: adaptado de Assumpção (2020, p. 135-136).

Dentre os campos apresentados no Quadro 4, destaca-se aqui o campo 375 de gênero que foi introduzido em 2009. Na definição de campo e escopo, está como “Informações sobre o gênero de uma pessoa, incluindo datas aplicáveis” (Library of Congress, 2022, não paginado, tradução nossa).

Os subcampos são: \$a (Gênero) indica o gênero com o qual uma pessoa se identifica, enquanto os subcampos \$s (Período inicial) e \$t (Período final) registram, respectivamente, a data de início e fim dessa identificação, caso aplicável. O subcampo \$u (Identificador Uniforme de Recursos - URI) associa um identificador externo ao registro, e o subcampo \$v (Fonte de informação) documenta a origem dos dados registrados. Para controle e vinculação, o subcampo \$0 (Número de controle de registro de autoridade ou número padrão) armazena um identificador único, enquanto o subcampo \$1 (URI de objeto do mundo real) fornece uma referência persistente.

Já o subcampo \$2 (Fonte do termo) especifica o código que identifica a fonte de um vocabulário controlado ou código para o gênero em \$. De acordo com a LC (2021, não paginado, tradução nossa), “Códigos de Origem de Gênero contém uma lista de código padrão e/ou fontes de termo para especificar o gênero de uma pessoa e atribui um código a cada fonte”. Alguns deles são, a ontologia de gênero, sexo, e orientação sexual (GSSO).

O subcampo \$6 (Ligação) permite a conexão entre diferentes versões linguísticas do registro, e o subcampo \$7 (Proveniência dos dados) registra a origem e confiabilidade das informações. Por fim, o subcampo \$8 (Link de campo e número de sequência) estabelece relações entre diferentes campos no registro bibliográfico (Library of Congress, 2022, não paginado, tradução nossa).

Por exemplo, o registro da autora Cecília Meireles utilizando o campo 375 ficaria conforme a seguir:

```
100 1# $a Meireles, Cecília, $d 1901-1964
370 ## $a Rio de Janeiro, Brasil $b Rio de Janeiro, Brasil $c
Brasil
374 ## $a Poeta $a Escritora $a Jornalista $a Professora
375 ## $a Feminino
377 ## $a por
400 1# $a Meireles, Cecília Benevides de Carvalho, $d 1901-1964
500 1# $a Pessoa associada: Manuel Bandeira (poeta)
670 ## $a Biblioteca Nacional do Brasil, Catálogo de Autoridades
$b (Cecília Meireles, 1901-1964, escritora e poetisa brasileira)
678 ## $a Cecília Meireles foi uma poeta, escritora, professora
e jornalista brasileira. Considerada uma das maiores escritoras
da literatura brasileira, teve papel importante no movimento
modernista, destacando-se por sua poesia lírica e intimista.
```

Além de o MARC 21 possuir o campo 375 para registrar o gênero da entidade, o UNIMARC para autoridades já incluía essa informação desde 2001, por meio do campo 120, que permite o registro do gênero associado à entidade descrita. Esse campo apresenta valores pré-definidos, como "a" para feminino, "b" para masculino, "c" para transgênero, "u" para desconhecido e "x" para não aplicável.

Em 2023, foi lançada a edição *on-line* do Manual do Formato UNIMARC para Autoridades, substituindo versões anteriores e incorporando atualizações. Inicialmente desenvolvido para registrar formas autorizadas de nomes, títulos e assuntos, o formato evoluiu com a adoção do IFLA-LRM, permitindo a identificação explícita de entidades bibliográficas e suas relações. A nova edição reformulou definições de campos e subcampos para alinhar-se a esse modelo. Com essa harmonização, o UNIMARC agora suporta tanto a catalogação baseada em relações entre entidades quanto o esquema tradicional de dados bibliográficos e de autoridade (IFLA, 2024).

Sendo assim, a presença de um campo que registre o gênero não é algo recente, mas uma prática que já estava sendo utilizada por outros padrões de metadados, como o UNIMARC. Para compreender como esse tipo de dado pode ser representado de forma estruturada, é necessário considerar os modelos conceituais e o RDA, que fornecem os princípios e diretrizes para a organização dos metadados nos contextos descritivos e de autoridade. Na próxima subseção, serão abordados os modelos conceituais e o RDA.

2.3 Modelos conceituais e o RDA Oficial e Original

O ambiente digital trouxe novas exigências para a descrição e recuperação da informação, o que demanda modelos mais flexíveis e adaptáveis à interoperabilidade entre sistemas e à conectividade da informação. Para garantir essa estruturação, diferentes modelos conceituais foram desenvolvidos ao longo do tempo, que servem como base para a descrição, recuperação e interoperabilidade da informação em sistemas bibliográficos, arquivísticos e museológicos.

As modelagens conceituais são “[...] processo de abstração de um sistema real ou proposto para um modelo. É quase exato que a modelagem conceitual é o aspecto mais importante de um projeto de simulação” (Robinson *et al.*, 2010, p. 03, tradução nossa).

Dentre os principais modelos aplicados à representação da informação em bibliotecas, destacam-se aqueles desenvolvidos pela IFLA, como o *Functional Requirements for Bibliographic Records* (FRBR), o *Functional Requirements for Authority Data* (FRAD) e o *Functional Requirements for Subject Authority Data* (FRSAD), posteriormente unificados, em 2017, no *IFLA Library Reference Model* (IFLA LRM).

O FRBR organiza o universo bibliográfico em entidades, atributos e relacionamentos, sua estrutura conceitual se baseia no modelo Entidade-Relacionamento proposto por Peter Chen em 1976. De acordo com o *IFLA Study Group on the Functional Requirements for Bibliographic Records*, o estudo possui dois objetivos,

[...] O primeiro é fornecer uma estrutura claramente definida e organizada para relacionar os dados registrados nos registros bibliográficos às necessidades dos usuários desses registros. O segundo objetivo é

recomendar um nível básico de funcionalidade para os registros criados pelas agências bibliográficas nacionais (*IFLA Study Group on the Functional Requirements for Bibliographic Records*, 1998, p. 07).

No FRBR, existem 10 entidades, com atributos específicos para cada e são divididos em três grupos, o primeiro grupo inclui obra, expressão, manifestação e item, o segundo grupo possui pessoa e entidade coletiva, e, por fim, o terceiro grupo com conceito, objeto, evento e lugar (IFLA, 1998).

Os usuários podem realizar quatro tarefas ao utilizar um catálogo ou conjunto de dados bibliográficos: encontrar, a fim de localizar um recurso com base em atributos como título, autor ou assunto. Após a localização, é necessário identificar o recurso para verificar se corresponde ao que se procura. Em seguida, selecionar entre diferentes edições e formatos, e obter o acesso ao recurso, seja por meio de empréstimo, compra ou acesso digital (IFLA, 1998).

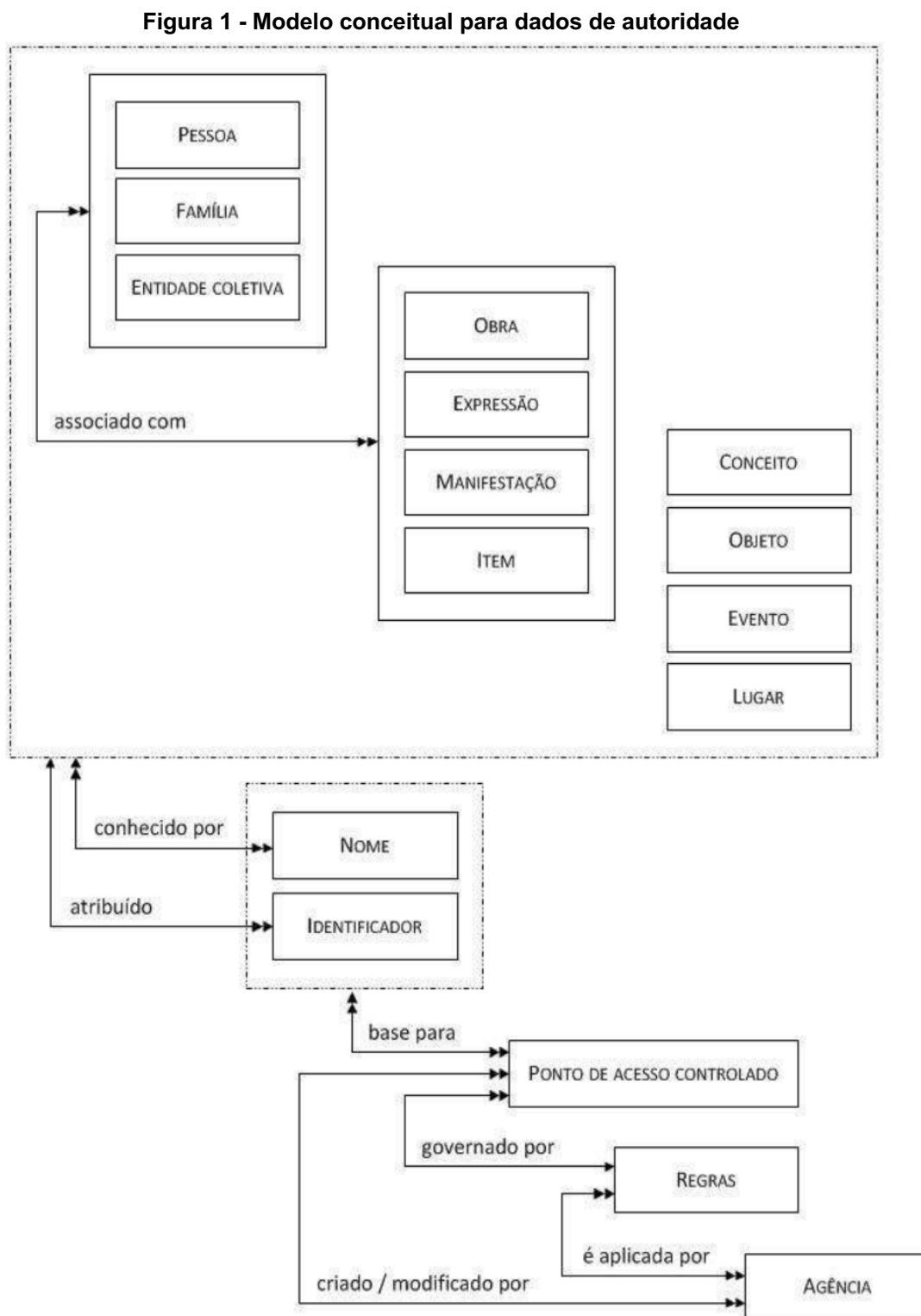
O FRBR forneceu uma nova visão sobre o universo bibliográfico e surge a necessidade de um modelo conceitual para os dados de autoridade. O FRAD foi desenvolvido para ampliar e detalhar as diretrizes voltadas à identificação dos elementos descritivos necessários para a criação e manutenção de dados de autoridade (Assumpção, 2012).

O *Working Group on Functional Requirements and Numbering of Authority Records* (FRANAR), definiu e adaptou as entidades e atributos com base no FRBR, UNIMARC para autoridades, *Mandatory Data Elements for Internationally Shared Resource Authority Records* (MLAR), *International Standard Archival Authority Record for Corporate Bodies, Persons and Families* (ISAAR(CPF)) e no *Guidelines for Authority Records and References* (GARR) (Assumpção, 2012).

Segundo a IFLA (2013, p. 1, tradução nossa), o objetivo é “[...] fornecer uma estrutura para a análise de requisitos funcionais para os tipos de dados de autoridade que são necessários para apoiar o controle de autoridade e para o compartilhamento internacional de dados de autoridade”.

Os relacionamentos do FRAD são organizados em quatro categorias distintas, aqueles que ocorrem entre as entidades do modelo em um nível mais genérico, os que conectam pessoas, famílias e entidades coletivas às obras, os que associam nomes a essas mesmas entidades, e, por fim, os que estabelecem vínculos entre os pontos de acesso controlados (Assumpção, 2012, p. 47).

Assumpção (2012), traduziu o modelo conceitual para dados de autoridade da IFLA, que está na Figura 1.



Fonte: Assumpção (2012, p. 48).

Para representar o modelo no contexto real, pode-se pensar os relacionamentos que a autora pode ter no modelo FRAD. Cecília Meireles pode ser representada como uma pessoa, associada às suas obras, como *Mulher ao Espelho*. Seu nome autorizado nos registros de autoridade é "Meireles, Cecília, 1901-1964", podendo incluir um identificador único, como um número VIAF. Suas obras são classificadas como uma obra, com expressões em diferentes edições e línguas, manifestações publicadas por diversas editoras e itens específicos em bibliotecas. O ponto de acesso controlado definido por regras, como o RDA ou AACR2, e mantido por instituições como a Biblioteca Nacional do Brasil. Além disso, outros conceitos podem ser relacionados, como objetos, representando manuscritos originais da autora, lugares, como o Brasil, onde atuou como escritora e educadora, e eventos, como o movimento modernista brasileiro, ao qual sua obra esteve ligada.

No modelo FRAD, o **gênero** é considerado um dos atributos da entidade pessoa. O modelo define atributos que permitem descrever uma pessoa em registros de autoridade, como nome, data de nascimento e morte, profissão, nacionalidade e gênero. No entanto, a inclusão do gênero nos registros de autoridade tem gerado debates na catalogação, principalmente pela abordagem binária (masculino/feminino), não conhecido e outro, e pela questão da privacidade.

Os usuários também possuem tarefas, a primeira é encontrar que permite localizar uma entidade ou um conjunto de entidades com base em atributos ou relacionamentos. A segunda identificar que confirma se a entidade recuperada corresponde à busca e distingue entre entidades semelhantes. Na sequência, contextualizar que esclarece relações entre pessoas, entidades coletivas e obras, situando-as em um contexto mais amplo. Por fim, justificar que documenta as escolhas na criação dos pontos de acesso controlados, garantindo transparência e rastreabilidade na catalogação (IFLA, 2013).

Para os dados de autoridade de assunto, foi criado o FRSAD. Esse modelo introduz duas entidades centrais, *Thema* e *Nomen*. O *Thema* é definido como "qualquer entidade utilizada como assunto de uma obra" (IFLA, 2010, p. 15), ou seja, representa o conceito ou tema tratado.

Já o *Nomen* corresponde a "qualquer signo ou sequência de signos (caracteres alfanuméricos, símbolos, sons, etc.) pelo qual um *thema* é conhecido, referenciado ou designado" (IFLA, 2010, p. 18). Dessa forma, o FRSAD possibilita uma abordagem

mais flexível e abrangente para a representação de assuntos nos catálogos e sistemas de recuperação da informação.

Possui tarefas de usuários assim como o FRBR e o FRAD, a primeira é encontrar, que permite localizar um ou mais assuntos e suas denominações com base nos critérios definidos pelo usuário, utilizando atributos e relacionamentos. A segunda é identificar, que auxilia na distinção entre assuntos ou denominações semelhantes, garantindo que a escolha feita corresponda ao que foi buscado. A terceira é escolher, possibilitando a seleção do assunto ou denominação mais adequada às necessidades do usuário. Por fim, explorar permite navegar pelas relações entre assuntos e denominações, contribuindo para uma melhor compreensão da estrutura conceitual e terminológica de um domínio específico (IFLA, 2010).

O IFLA-LRM foi desenvolvido com o objetivo de unificar, consolidar e corrigir as inconsistências dos modelos anteriores de requisitos funcionais para registros bibliográficos e de autoridade da família FR, além de buscar uma abordagem mais compatível com os princípios do *Linked Data*. De acordo com Riva, Bœuf e Žumer (2017, p. 09) 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”.

Foi estabelecida uma estrutura hierárquica organizada em níveis, introduzindo subclasses, atributos e relacionamentos para aprimorar a modelagem das entidades bibliográficas. Uma das principais mudanças foi a criação da entidade *Res*, que está no nível mais elevado da hierarquia e definida como "qualquer entidade no universo do discurso" (Riva; Bœuf; Žumer, 2017, p. 21). Essa reformulação substituiu o conceito de *Thema*, originalmente introduzido no FRASAD, que foi renomeado para *Res*, ampliando sua aplicabilidade e eliminando a restrição de que apenas determinadas entidades poderiam atuar como assunto de uma obra (Arakaki, 2020, p. 173).

Foram mantidas as entidades Obra, Expressão, Manifestação e Item, originadas no FRBR, mas com ajustes nas definições e notas de escopo. Uma das principais mudanças do LRM foi a introdução da entidade Agente como uma superclasse, unificando as categorias Pessoa (do FRAD) e Agente Coletivo (que no FRBR e FRAD pertenciam ao Grupo 2). Além disso, o LRM incorporou aspectos temporais e espaciais, por meio das entidades Local e Intervalo de Tempo, ampliando

as possibilidades de representação de contexto nos registros bibliográficos (Arakaki, 2020, p. 173).

Os atributos são definidos como instâncias específicas das entidades e podem ser registrados conforme a necessidade de cada aplicação. Os relacionamentos existentes no LRM foram aprimorados e expandidos, permitindo uma abordagem mais flexível e abrangente (Riva; Bœuf; Žumer, 2017).

Assim como os demais modelos da família FR, o LRM define um conjunto de tarefas do usuário. A primeira tarefa é encontrar informações a partir de qualquer critério relevante, permitindo a localização de entidades nos registros. Em seguida, o usuário deve identificar os recursos, diferenciando-os de outros semelhantes. A terceira tarefa consiste em selecionar o recurso mais adequado, aceitando ou rejeitando as opções disponíveis conforme sua necessidade. A quarta é obter acesso ao recurso selecionado, garantindo sua disponibilidade para consulta ou uso. Por fim, a quinta tarefa permite explorar os recursos, utilizando os relacionamentos entre as entidades para navegar pelas conexões e ampliar a descoberta da informação (Riva; Bœuf; Žumer, 2017, p. 16).

Possemato (2025) aponta que os modelos conceituais transformaram a forma de catalogar, agora o foco se torna a identificação e descrição das entidades de acordo com os diferentes cenários em que podem ser representadas, deixando a catalogação mais dinâmica e com a representação mais detalhada.

Nesse contexto, o RDA em sua versão Original, possui orientações para a criação de metadados de recursos informacionais e foi criado para substituir o AACR2r. Publicado pela primeira vez em 1967, o AACR foi desenvolvido com foco em descrições bibliográficas voltadas para registros em papel (Assumpção, 2012, p. 51).

No entanto, ocorreram transformações na catalogação para se adequar aos ambientes informacionais, surgiu a necessidade de maior flexibilidade na representação da informação. Diante desse cenário, era preciso um novo código que atendesse às exigências dos ambientes digitais, permitindo uma estrutura mais adaptável e interoperável. Tillett (2011, tradução nossa), destaca que

O novo código de identificação de recursos é o resultado de muitos anos de colaborações internacionais; ele produz metadados estruturados e interconectados para o ambiente digital, garantindo que as bibliotecas continuem sendo relevantes, mesmo como um controle de autoridade na era da *web* semântica.

O RDA Original se baseia nos modelos FRBR, FRAD, FRSAD e algumas características do AACR2r trazendo flexibilidade e detalhamento, permitindo uma melhor descrição de recursos digitais e integração com a *Web Semântica* e *Linked Data*. Oliver (2021), ressalta que o RDA Original foi concebido para ser utilizado em múltiplos contextos, não se limitando aos sistemas tradicionais de gerenciamento de bibliotecas. Entre os objetivos centrais de sua criação, esteve o desenvolvimento de um padrão de metadados compatível com o *Linked Data* e que as bibliotecas deveriam aproveitar as tecnologias emergentes da *Web Semântica*.

O FRBR e FRAD, se organizam a partir da definição de entidades, atributos e relacionamentos e o RDA Original por ter como base esses modelos, sua estrutura foi dividida em duas partes. A primeira trata do registro dos atributos das entidades, enquanto a segunda aborda os relacionamentos entre elas. As quatro primeiras seções correspondem à descrição das entidades, enquanto as seis seguintes lidam com as conexões entre elas. Além disso, o RDA Original mantém a organização dos Grupos 1, 2 e 3 do FRBR, garantindo uma abordagem estruturada e alinhada ao modelo conceitual (Assumpção, 2012, p. 52).

O Quadro 5 apresenta o mapeamento entre alguns atributos da entidade *Pessoa* nos modelos FRAD e RDA Original, associado às tarefas do usuário (encontrar, identificar, contextualizar e justificar). Esse mapeamento evidencia a função de cada atributo na construção dos registros de autoridade e destaca a presença do gênero em ambos.

Quadro 5 - RDA/FRAD Mapeamento de atributos da entidade Pessoa

FRAD – Atributo	Instrução RDA Original	Tarefas do usuário
Conhecido por	9.2 – Nome da pessoa 9.2.2 – Forma preferida do nome da pessoa 9.2.3 – Forma variante do nome da pessoa	<ul style="list-style-type: none"> ● Encontrar ● Identificar
Datas associadas à pessoa	9.3 – Datas associadas à pessoa (data de nascimento, morte, atividade)	<ul style="list-style-type: none"> ● Identificar ● Contextualizar
Forma completa do nome	9.5 – Forma completa do nome	<ul style="list-style-type: none"> ● Identificar
Gênero	9.7 – Gênero	<ul style="list-style-type: none"> ● Contextualizar ● Justificar
Local de nascimento	9.8 – Local de nascimento	<ul style="list-style-type: none"> ● Contextualizar

Local de falecimento	9.9 – Local de falecimento	● Contextualizar
País	9.10 – País associado à pessoa	● Contextualizar
Idioma da pessoa	9.14 – Idioma da pessoa	● Contextualizar
Campo de atividade	9.15 – Campo de atividade	● Contextualizar
Profissão/Ocupação	9.16 – Profissão ou Ocupação	● Contextualizar
Biografia/História	9.17 – Informação biográfica	● Contextualizar ● Justificar

Fonte: adaptado de Assumpção (2020, p.133-135).

Entre os atributos listados, destacam-se: nome completo, gênero, datas associadas, local de nascimento e falecimento, país, idioma, campo de atividade, profissão/ocupação e biografia/história. Cada atributo está associado a uma orientação específica do RDA Original o que permite sua aplicação padronizada no registro de autoridade.

O quadro 6 também aponta que o atributo gênero, tratado na instrução 9.7 do RDA Original, cumpre funções importantes para *contextualizar* e *justificar* a identidade da pessoa registrada. Isso reforça sua relevância na construção de registros mais completos e significativos para os usuários dos catálogos. Na sequência, o comparativo entre os atributos do FRAD, RDA e MARC 21 de autoridade.

Quadro 6 - Comparativo entre FRAD, RDA e MARC 21 de autoridade

FRAD - Atributo	Instrução RDA	MARC 21 Autoridade
Não contemplado	9.2 – Nome da pessoa	100 1# \$a / 100 2# \$a
Data associada à pessoa	9.3 – Datas associadas à pessoa	100 \$d
Gênero	9.7 – Gênero	375 \$a
Local de nascimento	9.8 – Local de nascimento	370 \$a
Local de falecimento	9.9 – Local de falecimento	370 \$b
País	9.10 – País associado à pessoa	370 \$c
Idioma da pessoa	9.14 – Idioma da pessoa	377 \$a
Campo de atividade	9.15 – Campo de atividade da pessoa	372 \$a
Profissão/Ocupação	9.16 – Profissão ou ocupação	374 \$a

Biografia/História	9.17 – Informação biográfica	678 \$a
--------------------	------------------------------	---------

Fonte: adaptado de Assumpção (2020, p. 141-142).

Conforme apresentado no Quadro 6, o atributo gênero está previsto no modelo FRAD e no RDA Original na instrução 9.7, como um elemento opcional, cuja aplicação depende das políticas e diretrizes estabelecidas por cada instituição catalogadora. Essa instrução prevê o uso de termos controlados como “masculino”, “feminino” e “desconhecido”, oferecendo uma padronização para o registro dessa informação em contextos bibliográficos e de autoridade. No formato MARC 21, o atributo gênero é registrado no campo 375 no subcampo \$a de autoridade.

A incorporação desse atributo nos registros de autoridade tem gerado debates técnicos e éticos, sobretudo no que se refere à privacidade das pessoas descritas, à exigência de consentimento prévio e à adequação do modelo para contemplar identidades de gênero não binárias. Esses pontos evidenciam a necessidade de revisão contínua das práticas de catalogação, de forma a garantir as representações de forma ética.

Como apontado anteriormente, os processos e instrumentos da catalogação seguem em constante transformação, com destaque para a publicação do IFLA LRM em 2017 e, posteriormente, a versão oficial do RDA em 2021, resultado do RDA *Toolkit Restructure and Redesign Project* (Projeto 3Rs). Essa atualização trouxe mudanças estruturais voltadas à clareza, flexibilidade e adequação ao ambiente digital, consolidando o RDA como um vocabulário de dados alinhado às práticas do *Linked Data*, o que promove a interoperabilidade e a sustentabilidade dos dados bibliográficos e de autoridade (Arakaki, A.; Arakaki, F., 2023).

Essa abordagem representa uma ruptura em relação a instrumentos anteriores, como o AACR2r e o MARC21, concebidos para contextos analógicos e com limitações quanto à legibilidade por máquinas. O RDA Oficial oferece ao catalogador um conjunto de elementos descritivos adaptáveis às necessidades dos usuários e aos diferentes tipos de recursos, além de possibilitar a conexão dos dados a ambientes digitais e à Web. Todos os seus elementos e vocabulários estão registrados em RDF no *Open Metadata Registry* (OMR), favorecendo a publicação e a reutilização semântica em diferentes contextos (Arakaki, A.; Arakaki, F., 2023). Na próxima subseção, serão discutidos os registros de autoridade, cuja estrutura e funcionamento refletem essas transformações e ampliam as possibilidades de representação das entidades.

2.4 Registro de autoridade

Os registros de autoridade estabelecem formas padronizadas de representação para nomes, entidades e conceitos. Taylor e Joudrey (2009, p. 252, tradução nossa) apontam que um registro de autoridade “inclui evidências de todas as decisões tomadas e de todos os relacionamentos entre as variações identificadas”. Sendo assim, os registros de autoridade favorecem a transparência e a rastreabilidade das informações.

As documentações das decisões e relações estabelecidas entre nomes e entidades permite um acompanhamento detalhado da construção e evolução dos registros ao longo do tempo, o que contribui para a análise da representatividade e da diversidade na catalogação, possibilitando que variantes e identificadores ampliem a visibilidade de grupos na produção científica e cultural.

A Declaração de Princípios Internacionais de Catalogação (PIC) da IFLA, publicada em 2016, trouxe uma mudança conceitual ao substituir o termo "registro" por "dados" nos contextos de autoridade e descrição bibliográfica. Essa alteração reflete uma abordagem mais flexível e alinhada às tecnologias atuais de gestão da informação, reconhecendo que o registro é apenas uma das formas possíveis de estruturar e apresentar os dados. Portanto, segundo o PIC (2016, p. 14) dados de autoridade são

Um conjunto de informações sobre uma pessoa, família, entidade coletiva, obra, expressão, manifestação, item ou thema. Os dados de autoridade devem ser elaborados para controlar as formas autorizadas do nome, nomen, formas variantes do nome e identificadores utilizados como pontos de acesso.

Os dados de autoridade possuem um papel amplo na documentação das decisões da catalogação e na rastreabilidade da informação. Nesse sentido, Assumpção (2018, p. 32) complementa essa visão ao afirmar que os dados de autoridade correspondem a

[...] o conjunto de dados formado pelos pontos de acesso controlados utilizados na identificação de uma determinada entidade e pelos dados que apoiam o uso destes pontos de acesso, sejam eles dados sobre a entidade, por exemplo, pontos de acesso relacionados, dados biográficos ou históricos, ou sobre o processo de estabelecimento dos pontos de acesso, por exemplo, as regras utilizadas, as fontes de informação consultadas e a instituição na qual ou para a qual os dados de autoridade foram criados.

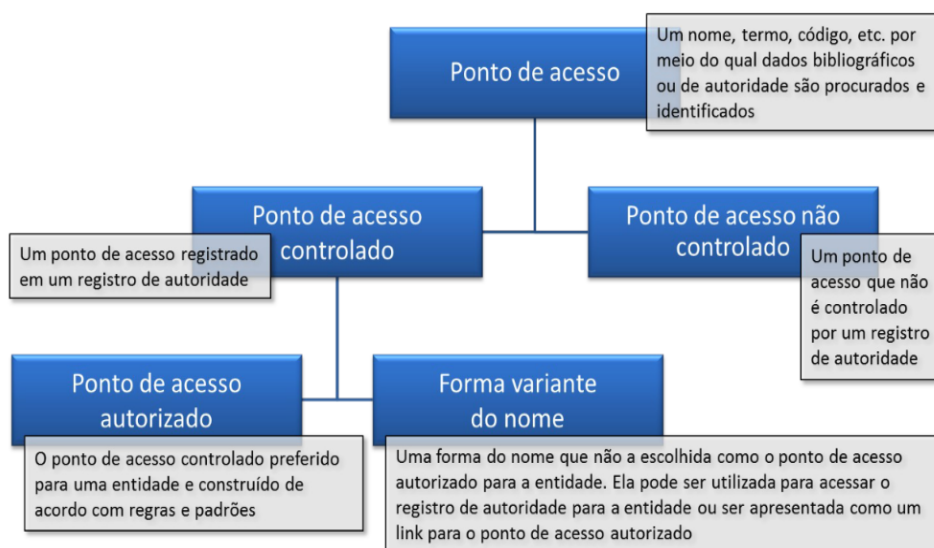
Os dados de autoridade fornecem as informações estruturadas sobre as entidades, e, para que os usuários acessem esses dados é necessário utilizar os

pontos de acesso, que são os elementos pelos quais se busca e recupera a informação em um sistema. Segundo Gorman (2004, p. 12, tradução nossa), o ponto de acesso

[...] leva o usuário ao registro; a descrição permite que o usuário decida se o item desejado é aquele que procura; e a localização conduz o usuário ao documento desejado. Essa é uma formulação simples e profunda, que serve como base para toda a catalogação.

De acordo com a PIC (2016), os pontos de acesso podem ser controlados ou não controlados, autorizados ou não, e são divididos de acordo com a entidade que representam, como ponto de acesso de nome pessoal, que se refere a uma pessoa, e ponto de acesso de assunto que representa um assunto. A Figura 2 a seguir elaborada por Assumpção (2012, não paginado), apresenta a estrutura dos pontos de acesso, destacando a distinção entre pontos de acesso controlados e não controlados, bem como a relação entre as formas autorizadas e variantes do nome.

Figura 2 - Pontos de acesso



Fonte: Assumpção (2012, não paginado).

A escolha dos pontos de acesso nos dados bibliográficos e de autoridade segue critérios específicos para garantir padronização e recuperação eficiente da informação. Nos dados bibliográficos, os pontos de acesso autorizados para obras e expressões são controlados para assegurar consistência na identificação da obra em diferentes contextos. Já o título da manifestação, geralmente não controlado, pode variar conforme a edição ou publicação do recurso. Além disso, os pontos de acesso

para criadores das obras são normalizados para agrupar corretamente suas produções (PIC, 2016).

Nos dados de autoridade, a padronização evita ambiguidades e possibilita um controle mais preciso da informação. Para isso, registra-se a forma autorizada do nome da entidade, estabelecendo um padrão de entrada conforme regras e diretrizes. Além disso, são incluídas formas variantes, abrangendo pseudônimos, nomes alternativos, traduções ou mudanças de grafia, garantindo que a entidade possa ser localizada independentemente da forma utilizada na busca (PIC, 2016).

Um exemplo prático dessa estrutura pode ser observado no registro da escritora Cora Coralina. O ponto de acesso autorizado segue a forma padronizada "Coralina, Cora, 1889-1985", garantindo consistência na recuperação da informação. Uma das formas variantes, registra-se "Seixas, Ana Lins dos Guimarães Peixoto Bretas, 1889-1985", correspondente ao seu nome de nascimento. Essa estruturação permite conectar diferentes formas do nome, promovendo precisão e acessibilidade nos registros catalográficos.

Os pontos de acesso autorizados são registrados como dados de autoridade, incorporando identificadores e formas variantes e servindo como referência padronizada nos catálogos. Para garantir consistência na recuperação da informação, os esquemas de descrição estabelecem e consideram a escolha do nome preferido para pessoas, famílias e entidades coletivas, além do título preferido para obras e expressões. A estruturação dos nomes respeita convenções culturais e linguísticas, diferenciando nomes semelhantes e registrando variantes para ampliar a acessibilidade e a precisão na identificação das entidades (IFLA, 2016).

O controle de autoridade envolve a padronização e gestão das formas de nomear e identificar pessoas, entidades e conceitos em um determinado sistema, ou seja, garante a consistência e unicidade dos pontos de acesso utilizados na catalogação e recuperação de informações. Gorman (2004, p. 21, tradução nossa) ressalta que

O controle de autoridade é central e vital às atividades que nós chamamos de catalogação. Catalogação – a reunião lógica de dados bibliográficos dentro de registros recuperáveis e utilizáveis – é aquela atividade que permite a biblioteca seguir sua missão central de servir acesso livre e aberto a todo conhecimento e informação registrada. Nós não podemos ter um serviço real de biblioteca sem uma arquitetura bibliográfica e nós não podemos ter essa arquitetura bibliográfica sem o controle de autoridade.

Assim, o controle de autoridade é utilizado por bibliotecários para padronizar nomes, títulos e assuntos em registros bibliográficos, garantindo consistência e clareza no acesso aos materiais. Ele assegura que variações de termos, como "filmes" e "cinema", sejam unificados sob um termo padrão, como "Filmes". Tillett (2004, p. 23) aponta que "O controle de autoridade é necessário para atender aos objetivos do catálogo, permitindo que os usuários encontrem as obras de um autor e agrupem todas as obras de uma pessoa ou entidade corporativa."

Dessa forma, o controle de autoridade garante a padronização das entradas nos catálogos, permitindo consistência, univocidade e confiabilidade na representação das entidades descritas, especialmente no que se refere à forma autorizada dos nomes e suas variantes. Além de assegurar a recuperação da informação, essa estrutura fornece a base para que os catálogos possam ser integrados a outros sistemas e repositórios, ampliando sua capacidade de uso e reuso em ambientes digitais expandidos.

A padronização de pontos de acesso e da descrição controlada de entidades, como pessoas, organizações e títulos, permite com que os registros possibilitem precisão na busca e na recuperação da informação. Contudo, a descrição baseada em formatos como o MARC 21 ou o UNIMARC, apresenta limitações no que diz respeito à interoperabilidade e à integração com outros sistemas e fontes de dados externos.

As transformações no ambiente digital e a necessidade de sistemas mais abertos, interconectados e capazes de lidar com grandes volumes de dados impulsionaram o desenvolvimento de novos paradigmas tecnológicos. Surgem as abordagens baseadas na modelagem de entidades e relacionamentos, que propõem a reorganização das informações bibliográficas em estruturas mais flexíveis e adaptáveis, como os modelos da família FR e o IFLA LRM. Essa mudança de perspectiva abre espaço para a adoção de tecnologias que permitem a descrição de dados de forma mais expressiva e relacional, superando as limitações.

A *Web Semântica*, concebida como uma extensão da Web atual, viabiliza a atribuição de significado explícito aos dados, permitindo sua interpretação tanto por humanos quanto por máquinas. A aplicação dos princípios da *Web Semântica* aos registros de autoridade potencializa o estabelecimento de conexões entre diferentes bases, o uso de identificadores persistentes, a integração com vocabulários controlados e o fortalecimento da interoperabilidade.

Dessa forma, os dados de autoridade e a Web Semântica se relacionam como elementos complementares, os dados de autoridade fornecem a estrutura descritiva e normativa, enquanto a *Web Semântica* oferece ferramentas tecnológicas para sua interligação, interoperabilidade e atualização contínua em ambientes digitais.

2.5 Catalogação e a inclusão de gênero

A catalogação, enquanto prática de representação da informação, tem evoluído para incorporar aspectos sociais, éticos e culturais na estruturação dos registros. Tradicionalmente voltada à normalização e ao controle de pontos de acesso, a catalogação passou a ser desafiada pela necessidade de representar a diversidade identitária dos sujeitos, sobretudo no que se refere à inclusão do gênero nos registros de autoridade. A construção de catálogos mais inclusivos implica reconhecer que os atributos atribuídos às entidades, como o gênero, não são neutros nem estáticos, mas socialmente construídos e historicamente situados.

Essa transformação está inserida em um cenário mais amplo de revisão ética e política das práticas catalográficas. Os sistemas de autoridade operam com base em uma ontologia estática e binária da identidade, o que desconsidera a fluidez, a complexidade e os contextos históricos que envolvem os sujeitos descritos.

Nesse sentido, a inclusão qualificada do metadado de gênero requer não apenas mudanças técnicas, mas também o fortalecimento de políticas baseadas na autodeclaração, no respeito à multiplicidade identitária e na rejeição de abordagens fixas ou essencialistas. A adoção de tecnologias da Web Semântica e de práticas baseadas em *Linked Data* tem ampliado as possibilidades de enriquecimento semântico dos registros, permitindo a vinculação a vocabulários controlados abertos e fontes externas confiáveis, como *Wikidata* e *VIAF*. Essas estratégias favorecem descrições mais granulares, contextuais e sensíveis à diversidade identitária.

A partir de 2016, esse debate foi incorporado de forma mais estruturada por organismos internacionais. O *Program for Cooperative Cataloging (PCC)* criou o *Ad Hoc Task Group on Gender in Name Authority Records* com o intuito de orientar o uso do campo 375 do MARC 21, prevendo os termos “masculino”, “feminino” ou “desconhecido”. No entanto, as críticas quanto à binariedade, à ausência de autodeclaração e à complexidade das identidades de gênero levaram à revisão

dessas diretrizes. Em 2022, o grupo passou a recomendar a não utilização do campo 375, enfatizando os riscos de violação de privacidade e reforço de vieses descritivos.

Apesar disso, outros padrões, como o UNIMARC, continuam prevendo campos específicos para o registro de gênero (campo 120), com opções mais amplas, como “feminino”, “masculino”, “transgênero”, “desconhecido” e “não aplicável”. Essa divergência evidencia a falta de consenso internacional e resulta em práticas institucionais heterogêneas, ora mais conservadoras, ora mais inclusivas.

A escassez de políticas institucionais claras e o receio de impor classificações fixas limitam a presença do metadado de gênero nos catálogos, impactando diretamente a visibilidade de autoras, pesquisadores e sujeitos cujas identidades não se encaixam em modelos normativos. Como argumenta Possemato (2025), a transição da catalogação para modelos baseados em entidades exige que atributos como o gênero sejam considerados em sua dimensão relacional, contextual e dinâmica.

Thompson (2016), por exemplo, analisou registros MARC 21 de autoridade para autores trans e observou que o uso do campo 375, presente em cerca de 70% dos registros, muitas vezes resultava em *outing* e reforçava uma lógica binária, desrespeitando a fluidez das identidades. A autora alerta para o risco de violação da autonomia dos sujeitos descritos, inclusive por meio de qualificadores e referências médicas. Como alternativa, a autora propõe que os autores tenham maior autonomia para se autoidentificar por meio de sistemas como o ORCID ou plataformas baseadas em *Linked Data*, que permitem a conexão e interoperabilidade entre diferentes conjuntos de dados. Ela defende uma abordagem de catalogação mais ética, pautada na privacidade, no consentimento e no respeito à autoidentificação, visando evitar práticas discriminatórias e promover uma representação mais justa nos registros de autoridade.

De modo semelhante, Billey, Drabinski e Roberto (2014) criticam a interpretação da regra 9.7 do RDA pela LC, que orientava a inclusão do gênero com base em suposições sobre nome ou aparência. Sob uma perspectiva da teoria queer, os autores argumentam que tais práticas impõem um modelo normativo excludente e desconsideram a complexidade histórica e social da identidade de gênero. Embora reconheçam que a ausência do gênero possa afetar a recuperação de obras, defendem que a categorização binária não deve ser assumida como necessidade do catálogo.

A esse respeito, Adolpho (2019) também aponta que a imposição de uma identidade fixa nos registros não contempla a diversidade de gênero e pode perpetuar formas de exclusão. Defende-se, portanto, a adoção de práticas baseadas na flexibilidade e no consentimento dos sujeitos representados.

A resposta institucional a essas críticas culminou, em 2022, na revisão das diretrizes do PCC, que recomendou não utilizar mais o campo 375 em novos registros e, nos casos de modificação, sua retirada. Informações sobre gênero passaram a ser permitidas apenas de forma não estruturada, nos campos 670 (notas) e 678 (notas biográficas).

Trans Metadata Collective et al. (2022) oferecem orientações para que instituições GLAM (Galerias, Bibliotecas, Arquivos e Museus) promovam práticas de metadados mais éticas e inclusivas. O documento destaca como os padrões atuais de catalogação frequentemente resultam na má representação ou no *outing* involuntário de pessoas trans, o que pode gerar riscos à privacidade e à segurança. Recomenda-se eliminar a menção ao gênero nos registros de autoridade (por exemplo, o campo 375 do MARC), salvo quando for essencial e baseado em consentimento informado. Também são incentivadas práticas como o uso da autodeclaração, a supressão de nomes anteriores quando solicitado, e a adoção de vocabulários inclusivos, como o *Homosaurus*. De forma geral, as diretrizes promovem justiça, autonomia e uma representação respeitosa para pessoas trans e de gêneros diversos.

Algumas instituições, como a Biblioteca Nacional da Suécia, também aderiram à diretriz do *International Standard Name Identifier* (ISNI), que optou por não incluir informações de gênero em seus registros. Essa orientação reforça a tendência de alinhamento com práticas mais respeitosas à autodeterminação.

Ao longo dessa discussão, destaca-se o papel dos catalogadores como agentes ativos na representação dos sujeitos. O Comitê Diretor de Ética de Catalogação reconhece essa responsabilidade e, por isso, publicou o *Cataloguing Code of Ethics* (2021), estabelecendo princípios éticos para decisões conscientes e justas na criação de registros de autoridade. Olson (2001, p. 4) lembra que, ao representarmos assuntos, influenciemos o acesso à informação e, conseqüentemente, ao conhecimento. Isto é, Olson (2001, p. 4) explica que

nossas teorias, modelos e descrições, aplicados no desempenho de nosso papel de intermediários entre as pessoas e a informação, são tão presunçosos e controladores quanto a construção e controle da natureza pelos cientistas. Enquanto bibliotecários, nós decidimos como representar os

assuntos e, assim, influenciaremos o acesso à (e o uso da) informação contida nos documentos que catalogamos e o acesso ao (e o uso do) conhecimento deles oriundo.

Como reforça Barbosa (1978, p. 30), a catalogação é o “[...] processo técnico do qual resulta o catálogo”, sendo este um instrumento de comunicação que só é eficaz se for normalizado. A IFLA (2016) também estabelece que os catálogos devem permitir que os usuários localizem, identifiquem, selecionem e obtenham recursos, inclusive por meio da navegação em registros bibliográficos e de autoridade.

Nesse contexto, a discussão sobre gênero nos registros de autoridade tem ganhado espaço entre autores como Billey, Drabinski e Roberto (2014), Thompson (2016), Adolpho (2019), Wainer (2019), Carbajal (2020), Martin (2021), Lindsey, Suiter e Hanselman (2022) e Possemato (2025), além dos documentos do PCC publicados em 2016 e 2022.

O RDA Original, por sua vez, trata do registro do gênero na instrução 9.7 da Seção 3, de forma não obrigatória, prevendo os termos “masculino”, “feminino” e “não informado”. No entanto, a orientação dada pela LC e pela NACO foi baseada em inferências que desconsideravam a autodeclaração (Martin, 2021).

A proposta de diretrizes apresentada nesta pesquisa dialoga com essas críticas e propõe o uso de vocabulários abertos, URIs e fontes confiáveis para enriquecer semanticamente os registros, respeitando a pluralidade identitária.

A revisão teórica desenvolvida nesta dissertação permitiu evidenciar os fundamentos técnicos e conceituais sobre representação da informação, metadados de autoridade e inclusão do gênero. Foram abordados os tipos e funções dos metadados, os principais padrões descritivos utilizados e os modelos conceituais da IFLA, com destaque para o IFLA LRM.

Constatou-se que, embora existam campos específicos para o registro do gênero (como os campos 375 do MARC 21 e 120 do UNIMARC), sua aplicação ainda é limitada. Isso revela um distanciamento entre o potencial técnico dos formatos e sua efetiva utilização nas bibliotecas. A ausência desse metadado compromete a visibilidade de sujeitos diversos e impacta diretamente os processos de recuperação e representação da informação.

Paralelamente, o uso de tecnologias de *Web Semântica* e *Linked Data* amplia as possibilidades de conexão e interoperabilidade entre sistemas, tornando viável uma representação mais contextualizada e respeitosa da identidade de gênero.

Com base nesse panorama, a pesquisa avança para a análise empírica das práticas adotadas por bibliotecas nacionais em relação à inclusão do metadado de gênero nos registros de autoridade, com o objetivo de propor diretrizes que articulem aspectos técnicos, sociais e tecnológicos, visando à construção de um sistema mais ético e inclusivo de representação da informação.

3 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa caracteriza-se como descritiva e exploratória, com abordagem qualitativa, voltada à compreensão das práticas de representação da identidade de gênero em registros de autoridade de bibliotecas nacionais. Os procedimentos técnicos adotados incluem pesquisa bibliográfica, documental e análise empírica de registros disponíveis online.

A pesquisa descritiva tem como objetivo principal a identificação e descrição das características de determinada população ou fenômeno, conforme Gil (2008). Já a pesquisa exploratória busca proporcionar maior familiaridade com o problema, tornando-o mais claro ou permitindo a formulação de hipóteses (Prodanov; Freitas, 2013). Conforme Gil (2008), esse tipo de abordagem permite examinar os temas sob múltiplos ângulos e é apropriada quando se trata de assuntos ainda pouco investigados. Nesse contexto, a pesquisa investigou a intersecção entre as seguintes temáticas: representação da informação, catalogação, metadados, gênero, dados de autoridade, controle de autoridade, *Web Semântica* e enriquecimento semântico.

Em relação à sua natureza, a pesquisa pode ser classificada como teórica considerando diretrizes e boas práticas estabelecidas por instituições como a *International Federation of Library Associations and Institutions (IFLA)* e o *World Wide Web Consortium (W3C)*.

É adotado abordagem qualitativa e, quanto aos procedimentos técnicos, configura-se como bibliográfica e documental. A pesquisa bibliográfica fundamenta-se na análise de materiais já publicados, como livros e artigos científicos. Por sua vez, a pesquisa documental recorre a fontes mais diversificadas e dispersas, como relatórios institucionais, normativas técnicas e documentos oficiais, que ainda não passaram por tratamento analítico aprofundado. Conforme Gil (2008), ambas as estratégias permitem conhecer e interpretar o estado atual do conhecimento sobre o tema investigado, contribuindo para a formulação de uma base teórica sólida.

Como critérios de inclusão, serão considerados artigos disponibilizados na íntegra, publicados em português, inglês ou espanhol. O recorte temporal abrange publicações entre 2000 e 2024, de modo a garantir a inclusão de estudos contemporâneos e alinhados à evolução do tema.

Para garantir a abrangência e qualidade dos estudos analisados, foram consultadas bases de dados nacionais e internacionais reconhecidas na área de

Ciência da Informação. No contexto nacional, destacam-se a Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI) e os periódicos especializados disponíveis no Portal de Periódicos da CAPES. No cenário internacional, a seleção considerou as bases *Library and Information Science Abstracts* (LISA), *Web of Science*, *Scopus* e *Scientific Electronic Library Online* (SciELO), possibilitando a incorporação de estudos consolidados e recentes sobre a temática.

Em relação ao enriquecimento semântico, houve a necessidade de aprofundar como as instituições LAM estão realizando o processo, assim, foi conduzida por meio de uma revisão da literatura, que, segundo Gil (2008), permite mapear o conhecimento já produzido sobre um tema, identificando tendências, abordagens e lacunas na área de estudo. Em relação ao objetivo, de acordo com Gil (2004, p. 42), "As pesquisas descritivas têm como objetivo primordial a descrição das características de determinada população ou fenômeno". Esse tipo de abordagem busca organizar, classificar e interpretar dados para apresentar um panorama detalhado da realidade investigada.

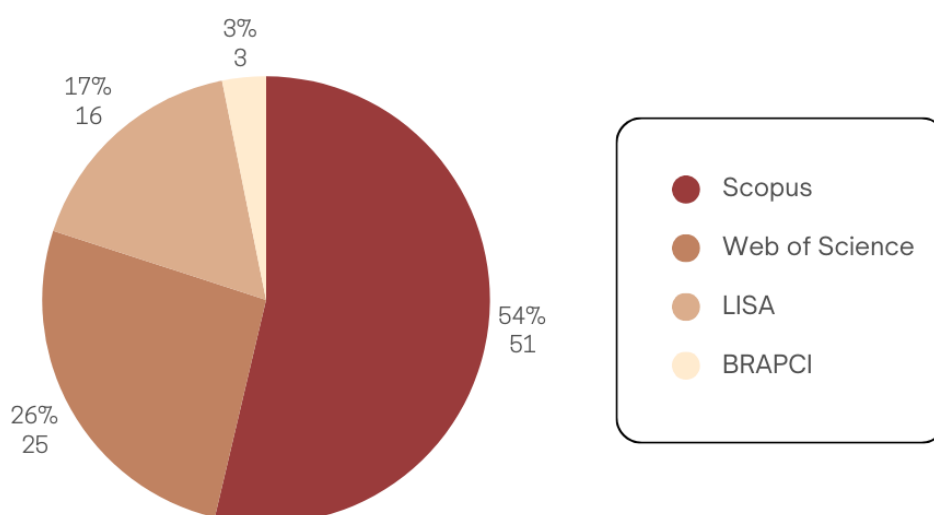
Para isso, foi utilizada a expressão de busca "*semantic enrichment*" AND "*metadata*", aplicada no dia 14 de fevereiro de 2025, em quatro bases de dados: *LISA*, *Scopus*, *Web of Science* (WoS) e BRAPCI. A base SciELO também foi consultada, porém, não apresentou resultados compatíveis com os critérios da pesquisa.

Tendo em vista a temática deste estudo, realizou-se uma busca a fim de verificar a existência de estudos que abordam especificamente a relação entre o enriquecimento semântico de metadados e o gênero. Foram combinados os termos "*semantic enrichment*" AND "*metadata*" AND "*gender*". No entanto, essa estratégia não recuperou documentos nas bases consultadas, evidenciando uma lacuna na literatura científica sobre o tema. Esse cenário reforça a necessidade de investigações que explorem o potencial do enriquecimento semântico para aprimorar a representação do gênero.

Foram considerados artigos publicados entre 2000 e 2024, disponíveis nos idiomas inglês, português e espanhol. O recorte temporal foi definido com o intuito de abranger os avanços na área ao longo das últimas duas décadas, acompanhando a evolução das pesquisas sobre metadados e semântica no contexto da Ciência da Informação.

A busca resultou na recuperação de 95 documentos, distribuídos entre as bases consultadas. A Figura 3 apresenta a distribuição dos 95 documentos recuperados nas bases consultadas. Na Scopus foram recuperados 51 documentos, representando 54% do total. Em seguida, na *Web of Science* (WoS) foram recuperados 25 documentos, correspondendo a 26%. A base LISA apresentou 16 documentos, o que equivale a 17% da amostra. Já a BRAPCI teve o menor volume de artigos indexados, totalizando três documentos, o que representa 3% do total recuperado.

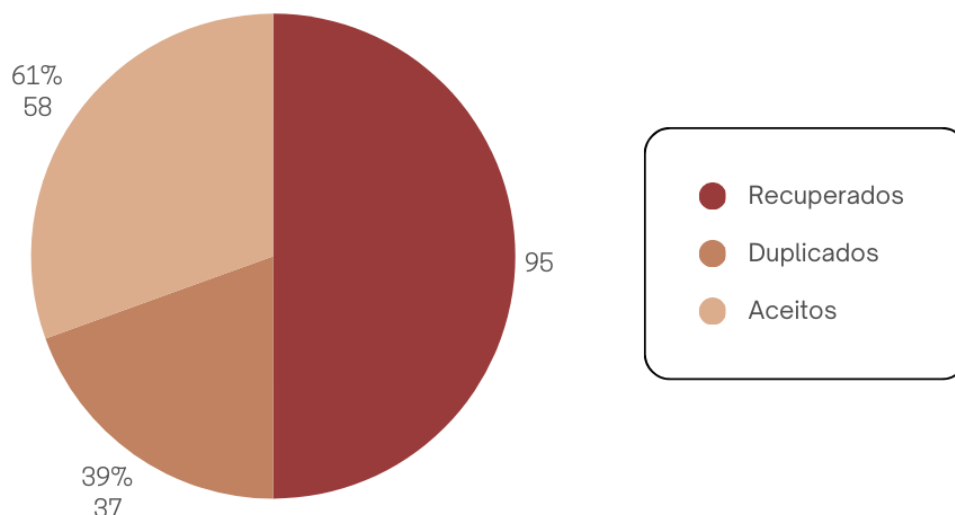
Figura 3 - Documentos recuperados por bases de dados



Fonte: Elaborado pela autora (2025).

Foi realizada uma etapa de refinamento, na qual foram eliminados 37 documentos duplicados, identificados em mais de uma base de dados. Essas duplicatas representavam 39% do total. Com a remoção, restaram 58 documentos, equivalentes a 61% da amostra.

Figura 4 - Documentos recuperados, aceitos e rejeitados



Fonte: Elaborado pela autora (2025).

A partir dos 58 documentos, foram selecionados apenas 19 artigos que tratam do enriquecimento semântico de metadados em Bibliotecas, Arquivos e Museus (LAM).

A pesquisa também pelo seu caráter descritivo e exploratório da pesquisa, foram identificadas bibliotecas nacionais das Américas e da Europa, utilizando como referência a *Wikipédia*³ que organiza essas instituições por continente. Foi realizada coleta de dados por meio da análise de registros de autoridade de pessoas nos catálogos das bibliotecas nacionais, em conjunto com o contato direto (e-mail) com essas instituições, com o objetivo de verificar a presença de informações sobre gênero e compreender as diretrizes relacionadas à sua representação.

Como parte da estratégia metodológica, foi realizado o envio de e-mails formais às instituições analisadas, com o objetivo de obter esclarecimentos sobre suas políticas e práticas relativas à inclusão de metadados de gênero em registros de autoridade. O contato buscou um posicionamento institucional, solicitado por meio de canais oficiais de comunicação, e não teve caráter de entrevista pessoal. As respostas foram tratadas como manifestações institucionais, representando as diretrizes e procedimentos adotados por cada biblioteca nacional no âmbito da gestão de seus registros de autoridade. Essa abordagem permitiu complementar a análise

³ As bibliotecas nacionais listadas pela Wikipédia, são: https://pt.wikipedia.org/wiki/Biblioteca_nacional

documental dos catálogos públicos com informações contextuais sobre as decisões e orientações internas referentes ao tratamento do dado de gênero.

A análise foi conduzida com base em uma abordagem comparativa e qualitativa, envolvendo:

- Identificação da ocorrência do metadado de gênero nos registros de autoridade de autoras selecionadas;
- Verificação do campo e subcampos utilizados, com ênfase nos campos 375 (MARC 21) ou 120 (UNIMARC), quando aplicável;
- Análise dos valores atribuídos ao metadado, com atenção à binariedade, uso de vocabulários controlados, presença de URIs e fontes de autoridade (como VIAF e Wikidata);
- Mapeamento de boas práticas e ausências de padronização, interoperabilidade e respeito à autodeclaração;
- Articulação entre os achados e a literatura científica previamente discutida, com foco nas implicações sociais, técnicas e éticas da presença (ou ausência) do metadado de gênero.

A seleção das bibliotecas nacionais analisadas nesta pesquisa baseou-se em um critério geográfico e institucional. Foram priorizadas bibliotecas nacionais localizadas nas Américas e na Europa, considerando a diversidade de contextos catalográficos, o acesso público aos registros de autoridade por meio de interfaces online e a relevância dessas instituições como centros de referência na produção e manutenção de dados bibliográficos e de autoridade. Como ponto de partida, utilizou-se a listagem de bibliotecas nacionais organizada pela *Wikipédia*, complementada por consultas aos portais institucionais de entidades vinculadas à IFLA. O critério central foi a disponibilidade de catálogos que permitissem a consulta direta a registros de autoridade de nomes pessoais, preferencialmente no formato MARC 21 ou com suporte à visualização de elementos estruturados.

Um dos desafios metodológicos foi a respeito às variações linguísticas encontradas nas interfaces dos catálogos e nos próprios registros, o que exigiu adaptações na terminologia utilizada nas buscas e na interpretação dos campos preenchidos. Em alguns casos, a visualização pública dos registros era limitada, dificultando o acesso completo aos dados de autoridade. Além disso, identificou-se que mesmo entre bibliotecas que utilizam o MARC 21, há diferenças na forma de

preenchimento dos campos e subcampos relacionados ao gênero, o que impacta na padronização dos dados e na comparabilidade entre os registros.

Apesar dessas limitações, os procedimentos adotados permitiram identificar um conjunto representativo de práticas e estratégias adotadas por diferentes bibliotecas nacionais, possibilitando uma análise crítica do uso do metadado de gênero e do potencial das diretrizes baseadas em enriquecimento semântico.

3 WEB SEMÂNTICA E LINKED DATA

Em 1989, Tim Berners-Lee, criou a *World Wide Web* (WWW), baseada em três tecnologias, o *Hypertext Markup Language* (HTML), o servidor HTTP (*Hypertext Transfer Protocol*) e o *Unified Resource Identifier* (URI), com a finalidade de auxiliar no compartilhamento de documentos entre pesquisadores do CERN permitindo o acesso e a troca de informações de maneira mais eficiente (Isotani; Bittencourt, 2015).

A primeira conferência internacional sobre a WWW ocorreu em 1994, durante a qual foi anunciado o *World Wide Web Consortium* (W3C), responsável por definir padrões e tecnologias para o desenvolvimento da *Web*. Na ocasião, Berners-Lee destacou que a ausência de semântica na web limitava sua funcionalidade, permitindo apenas a leitura por humanos, sem possibilitar que máquinas interpretassem e relacionassem os dados de maneira autônoma (Isotani; Bittencourt, 2015).

Essa limitação restringia a busca, recuperação e reutilização da informação, além de dificultar a interoperabilidade entre sistemas, já que os dados não possuíam estrutura semântica para estabelecer conexões automatizadas (Berners-Lee, Hendler, Lassila, 2001; Hitzler, Krötzsch, Rudolph, 2010).

Para resolver esse problema, Berners-Lee, Hendler e Lassila (2001) propuseram a *Web Semântica*, que atribui significado explícito aos dados, permitindo que sejam interpretados e processados por máquinas. Diferente da *Web* convencional, onde a compreensão da informação depende exclusivamente dos usuários, a *Web Semântica* organiza os dados de modo que computadores e pessoas possam interagir de maneira integrada.

Berners-Lee, Hendler e Lassila (2001, p. 2, tradução nossa) apontam que a *Web Semântica* é uma extensão da *Web* atual “[...] na qual as informações recebem um significado bem definido, permitindo que os computadores e as pessoas trabalhem em cooperação”.

Foi desenvolvida com base em um modelo em camadas, o chamado “bolo de noiva”, integrando diversas tecnologias e padrões (Isotani; Bittencourt, 2015). Entre esses componentes estão ontologias, que formalizam o conhecimento e definem relações entre conceitos, consultas semânticas, como as realizadas por meio do

Simple Protocol and RDF Query Language (SPARQL); o *Resource Description Framework* (RDF) utilizado para modelar e interligar dados; e a *Ontology Web Language* (OWL), que permite a criação de vocabulários controlados para representar informações de maneira mais sofisticada e interoperável.

Ainda no desenvolvimento da *Web Semântica*, em 2006, Berners-Lee introduziu o conceito de *Linked Data*, definido como um conjunto de práticas para publicação e interligação de dados estruturados na *Web*, permitindo que sejam recuperados e interpretados automaticamente por máquinas (Berners-Lee, 2006).

De acordo com Bizer, Heath e Berners-Lee (2009, não paginado, tradução nossa), o *Linked Data*

refere-se a dados publicados na *Web* de tal forma que sejam legíveis por máquina, seu significado seja explicitamente definido e sejam ligados a outros conjuntos de dados externos que, por sua vez, possam ser vinculados a partir de conjuntos de dados externos.

Para Pomerantz (2015, p. 157, tradução nossa), “A *Web Semântica* exige mais do que simplesmente disponibilizar dados online, é necessário estabelecer conexões entre dados estruturados”. Para a estruturação dos dados em *Linked Data* Berners-Lee (2006, não paginado, tradução nossa) destaca quatro princípios

1. Use URIs como nome para coisas;
2. Use URIs HTTP para que as pessoas possam procurar esses nomes;
3. Quando alguém procurar um URI, forneça informações úteis usando os padrões RDF e SPARQL;
4. Inclua *links* para outros URIs para que elas possam descobrir mais coisas.

O primeiro princípio é a utilização de URIs como nome para coisas, segundo Jesus, Castro e Ramalho (2021, p. 6) os URIs são [...] “identificadores que servem para representar documentos, locais, endereços da *Web*, figuras ou até mesmo pessoas”. Um exemplo do uso de URIs em registros bibliográficos e de autoridade é no *Virtual International Authority File* (VIAF), que associa diferentes registros de autoridade a um identificador único. Por exemplo, a escritora Cecília Meireles possui

um URI⁴ específico no VIAF, que permite com que as instituições recuperem de forma padronizada as informações, independente das variações.

O segundo princípio é a utilização de URIs HTTP, para identificar objetos e conceitos abstratos, permitindo que esses identificadores sejam acessíveis e referenciados na *Web* (Heath; Bizer, 2011). O terceiro princípio enfatiza o uso dos padrões RDF e SPARQL. De acordo com a W3C (2014, não paginado, tradução nossa), o RDF “[...] fornece uma estrutura comum para expressar essas informações, de modo que elas possam ser trocadas entre aplicativos sem perda de significado”. Com ele é possível estabelecer conexões entre diferentes fontes de informação, permitindo o acesso integrado a dados distribuídos. Essa abordagem diferencia a *Web* de Dados da *Web* de Documentos, pois possibilita a interoperabilidade e links estruturados (Isotani; Bittencourt, 2015).

O SPARQL além de ser uma linguagem de consulta é “[...] um protocolo para recuperação e manipulação de dados armazenados no formato *Resource Description Framework* (RDF). Ele fornece um meio para consultar diversas fontes de dados, sejam elas locais ou distribuídas na *Web*” (W3C, 2013, não paginado, tradução nossa).

Por fim, o quarto princípio é a inclusão de *links* para outros URIs, permitindo a conexão entre diferentes fontes de dados. Por exemplo, no *Wikidata*, utiliza-se URIs para interligar informações sobre pessoas, organizações, obras e conceitos a outras bases de conhecimento, como a LC e a Biblioteca Nacional da França (BNF). Dessa forma, ao acessar um item no *Wikidata*, é possível encontrar links diretos para registros de autoridade dessas instituições, promovendo maior interoperabilidade e enriquecimento da informação.

Baker *et al.* (2011, não paginado, tradução nossa) apresentam a seguinte definição

São considerados dados em *Linked Data*, dados conectados, aqueles publicados de acordo com os princípios definidos para facilitar o vínculo entre *datasets*, ou seja, utilizando URLs como identificadores únicos de recursos, expressos sob o modelo RDF, especificando as relações existentes entre as coisas.

⁴ URI Cecília Meireles no VIAF: <https://viaf.org/en/viaf/7405447>.

Ou seja, para atingir o propósito do desenvolvimento da *Web* de dados, é essencial aderir aos padrões recomendados pelo W3C.

Observa-se que o *Linked Data*, por objetivar que os conjuntos de dados sejam interligados e descritos de forma correta, relacionando-se com o conceito de dados abertos, cria o conceito do *Linked Open Data* (LOD), no qual os documentos são disponibilizados de forma aberta e a todos. De acordo com Arakaki (2016, p. 22)

[...] no ambiente digital, os dados abertos (*Open Data*) estão cada vez mais acessíveis à consulta comum e, ao mesmo tempo, à ligação entre eles. Assim, os dados estarão mais suscetíveis à ligação de dados abertos (*Linked Open Data*) na *Web*, facilitando a interoperabilidade e o acesso à informação.

No contexto de dados e conteúdos abertos, firma-se o significado de que qualquer pessoa pode acessar, usar, modificar e compartilhar livremente para qualquer finalidade, sujeito, no máximo, aos requisitos que preservam a procedência e a abertura (*Open Knowledge Foundation*, 2010, não paginado, tradução nossa).

Berners-Lee (2006) propõe uma classificação em estrelas para indicar o grau de abertura dos dados, onde um maior número de estrelas representa um nível mais elevado de abertura, conforme demonstrado no Quadro 7.

Quadro 7 - Classificação dados abertos

Classificação	Descrição
☆	Disponível na <i>web</i> em qualquer formato, mas com uma licença aberta, para ser Open Data
☆☆	Disponível como dados estruturados legíveis por máquina (por exemplo, Excel em vez de digitalização de uma imagem de uma tabela)
☆☆☆	Disponível como dados estruturados legíveis por máquina em formato não proprietário (por exemplo, CSV)
☆☆☆☆	Além das 3 primeiras estrelas, utilizar padrões abertos da W3C como RDF e SPARQL para identificar as coisas, de modo que as pessoas possam relacionar seu material
☆☆☆☆☆	Todos os demais e vincular seus dados para outros de pessoas para fornecer contexto

Fonte: adaptado de 5 Stars Open Data (2015).

A classificação dos dados abertos fornece um guia para a disponibilização e uso de dados estruturados na *Web*. Essa classificação, que varia de uma a cinco estrelas, representa diferentes níveis de acessibilidade, interoperabilidade e

reutilização dos dados. Diversas instituições como Facebook, Best Buy, Governo Britânico estão conectando os dados (Isotani; Bittencourt, 2015). Além dessas, instituições LAM já implementam as diretrizes em seus projetos para ampliar a conectividade e a descoberta de informações, por exemplo, a LC, a Europeia, Biblioteca Nacional da Alemanha (DNB), Biblioteca Nacional da França (BNF) e a Biblioteca Nacional de Espanha (BNE) que publica os catálogos bibliográficos e arquivos de autoridade em RDF e SPARQL.

O W3C estabeleceu 35 boas práticas para dados na *Web*. As "Boas Práticas para Dados na *Web*" enfatizam a importância de tornar os dados acessíveis, interoperáveis e reutilizáveis, com foco em padrões abertos e metadados de qualidade. O documento detalha recomendações para a publicação e gestão de dados na web, destacando a transparência, a preservação e a usabilidade dos dados. Ele também aborda questões éticas e de privacidade, promovendo a confiança e a inovação a partir de dados bem estruturados e documentados, sempre considerando a sustentabilidade no compartilhamento de informações na *Web*.

Silva (2013, p. 94) discorre sobre a diferença entre o *Linked Data* e *Linked Open Data*

Tendo em vista a necessidade de dados abertos para a implementação do projeto *Linked Data*, foi criado o *Linked Open Data*, que é justamente a ideia do *Linked Data* com enfoque na abertura dos (dados disponíveis na *Web*, para que possam ser efetivamente interligados.

Relacionando o *Linked Open Data* e a *Web Semântica*, Santarém Segundo (2015, p. 225) assim se manifesta

O *Linked Open Data*, que atualmente apresenta-se como a melhor forma de materialização dos conceitos e tecnologias da *Web Semântica*, é um projeto, com um conjunto de normas a serem seguidas, que usa os mesmos princípios de ligação semântica da *Web* de dados; entretanto, apresenta particularidades específicas, indicando um grau de exigência maior na constituição de sua rede de interligações.

Na prática, a DBpedia e Europeia exemplificam essa interconexão, permitindo que informações provenientes de diferentes fontes possam ser acessadas e reutilizadas de maneira integrada.

Além disso, a W3C (2014) definiu dez boas práticas para a publicação de LOD, são elas, o primeiro é preparar os stakeholders, explicando o processo de criação e manutenção do LOD. Em seguida, é necessário selecionar um conjunto de dados que seja relevante e possa ser reutilizado. O terceiro passo consiste em modelar os dados,

representando objetos e suas relações de forma independente de aplicações específicas. Após isso, deve-se especificar uma licença apropriada, garantindo que o reuso seja facilitado com uma declaração clara sobre origem, propriedade e condições de uso dos dados publicados.

O quinto passo é a definição de URIs adequadas para os dados vinculados, considerando uma estratégia de nomenclatura bem planejada e baseada em HTTP URIs, com suporte a múltiplos idiomas, controle de mudanças ao longo do tempo e garantia de persistência dos identificadores. O sexto passo é utilizar vocabulários padronizados, descrevendo os objetos com vocabulários já estabelecidos sempre que possível, estendendo-os conforme necessário e criando novos apenas quando indispensável.

A conversão dos dados para um formato compatível com o LD, normalmente realizada por meio de scripts ou processos automatizados. Além disso, deve-se fornecer acesso automatizado aos dados, permitindo que mecanismos de busca e outras aplicações os processem utilizando padrões da Web.

Após a conversão e disponibilização, é fundamental anunciar os novos conjuntos de dados em domínios autorizados, estabelecendo um compromisso social implícito com a comunidade que utilizará esses dados. Por fim, é necessário reconhecer essa responsabilidade social, garantindo a manutenção e a disponibilidade contínua dos dados conforme anunciado, assegurando sua confiabilidade ao longo do tempo.

Wood, Zaidman e Ruth (2014) e Baker (2012) enfatizam a importância das diretrizes e boas práticas fornecidas pela W3C para a implementação eficaz da *Web Semântica*. Com o avanço das novas tecnologias e o surgimento de ferramentas, essas diretrizes viabilizam a publicação em *Linked Data*.

Diante do exposto, observa-se que a publicação de LD e LOD seguem um conjunto estruturado de boas práticas que visam garantir a interoperabilidade, acessibilidade e reutilização dos dados. A implementação dessas diretrizes permite a conexão dos dados na *Web Semântica*, permitindo que informações sejam recuperadas e reutilizadas em diferentes contextos.

A *Web Semântica* transformou a forma como as instituições LAM (Bibliotecas, Arquivos e Museus) disponibilizam seus dados. A adoção de tecnologias da *Web Semântica*, boas práticas e diretrizes possibilitam que essas instituições estruturem

seus registros de maneira mais interconectada, permitindo uma recuperação de informação e ampliando o acesso a diferentes bases de conhecimento.

O próximo tópico discute sobre o enriquecimento semântico de dados LAM, destacando como essas instituições realizam o processo e utilizam ferramentas para aprimorar a estruturação, granularidade, descobertas e a conectividade dos seus registros.

3.1 Enriquecimento de metadados LAM

O termo "enriquecimento" refere-se ao processo de utilizar ferramentas que possibilitam a criação de novos metadados ou a melhoria dos já existentes, atribuindo-lhes significados mais contextualizados (Isaac *et al.*, 2015). Essa prática transforma a forma como os dados são tratados, tornando-os mais compreensíveis, interoperáveis e úteis. Com isso, abre-se caminho para sua conexão com outras bases de dados, viabilizando o acesso a informações que anteriormente estavam isoladas ou desconectadas.

À medida que as demandas informacionais se tornaram mais complexas, a necessidade de estratégias de enriquecimento tornou-se evidente. Peng *et al.* (2015) observam que metadados básicos, como data de publicação, autor e idioma, embora fundamentais, muitas vezes não são suficientes para atender às necessidades de pesquisadores. Por exemplo, em estudos sobre estilos de escrita de autores do século XIX, a ausência de informações sobre o gênero dos autores nos metadados convencionais dificulta a categorização dos dados, evidenciando a importância de se adicionar novas camadas de significado e contexto aos registros bibliográficos.

Nesse sentido, o enriquecimento de metadados não se resume à simples adição de informações, mas envolve a construção de relações semânticas entre os dados, ampliando a compreensão do recurso descrito. Como destacam Manguinhas *et al.* (2016, p. 01, tradução nossa), esse processo “[...] coloca os objetos em contexto, vinculando-os a entidades relevantes (pessoas, lugares, tipos de objetos etc.)”. Ao estabelecer essas conexões, o enriquecimento semântico contribui significativamente para a recuperação mais eficiente e contextualizada da informação.

No entanto, apesar das potencialidades atribuídas ao uso de URIs e vocabulários controlados externos em ambientes semânticos, a aplicação prática

dessas estratégias em registros MARC 21 de autoridade ainda enfrenta uma série de desafios. A estrutura legada do MARC, concebida para sistemas bibliográficos centralizados e pouco orientados à web, apresenta limitações importantes quanto à flexibilidade necessária para incorporar identificadores persistentes e múltiplos pontos de ligação com bases externas. Mesmo com a introdução de campos e subcampos como o 375 (Gênero), \$0 (número de controle de autoridade), \$1 (URI de objeto do mundo real) e \$u (link externo), a efetividade dessa vinculação depende da existência de políticas institucionais claras, de infraestrutura tecnológica adequada e da capacitação dos catalogadores na seleção criteriosa de fontes confiáveis.

Além dessas limitações, a ausência de interoperabilidade nativa entre o formato MARC 21 e modelos orientados a grafos, como o RDF, impõe a necessidade de camadas adicionais de conversão e mapeamento, frequentemente marcadas por inconsistências ou falta de transparência. Tais obstáculos técnicos e operacionais comprometem a consolidação de práticas de enriquecimento semântico, sobretudo em contextos institucionais onde a atualização tecnológica e a formação continuada ainda são incipientes.

Nesse panorama, Van Veen, Lonij e Koppelaar (2015, tradução nossa, não paginado) definem o enriquecimento semântico como a atribuição de significado aos dados por meio da inclusão de informações complementares. Essas adições podem assumir diferentes formas, desde a extração de características específicas como análise de sentimentos até a incorporação de dados associados, como coordenadas geográficas.

Com base nesse entendimento, diversos métodos e abordagens têm sido desenvolvidos para permitir que dados estruturados, semi-estruturados e não estruturados sejam enriquecidos de modo a aprimorar sua qualidade, ampliar sua capacidade de descoberta e reforçar seu valor informacional em ambientes digitais conectados, de acordo com Zeng (2019, p. 02, tradução nossa), alguns são

[...] a utilização de vocabulários de sistemas de organização do conhecimento (KOS) e outros recursos que adotaram Linked Open Data (LOD); transformar dados semi-estruturados em dados estruturados por meio de análise semântica baseada em entidades e anotação para ampliar os pontos de acesso; explorar dados não estruturados e gerar bases de conhecimento para apoiar a descoberta de conhecimento; possibilitar o uso de dados de LAM (Bibliotecas, Arquivos e Museus) em vários contextos, em silos de dados, ao mesmo tempo em que oferece interfaces de usuário intuitivas online; e tornar os conteúdos heterogêneos de diversos provedores semanticamente interoperáveis por meio de uma infraestrutura de ontologia compartilhada.

Isaac *et al.* (2015) pontua que existem muitas ferramentas que podem auxiliar no processo, porém elaborar uma lista completa é impossível já que é um campo em constante evolução. A variedade de abordagens e técnicas evidencia tanto a complexidade do processo em relação aos diferentes tipos de dados e suas aplicações quanto a sua flexibilidade, permitindo sua adaptação a diversos contextos e necessidades informacionais.

Um dos métodos destacados por Zeng (2019) é a utilização de recursos contextuais externos para o enriquecimento dos dados estruturados. De acordo com Smith-Yoshimura (2018), as fontes de dados vinculados mais bem classificadas segundo o estudo realizado pela OCLC em 2018 são do primeiro ao quinto: *id.loc.gov*; *VIAF*; *DBpedia*; *GeoNames* e *Wikidata*. Em comparação com a pesquisa em 2015, a mudança de maior destaque foi o aumento na utilização do *Wikidata*, que estava na 15 posição e foi elevado para a quinta posição.

No contexto desta pesquisa, é relevante destacar que tanto o *VIAF* quanto o *Wikidata* incluem a informação de gênero em seus registros, o que possibilita sua utilização como fontes externas para enriquecimento semântico. No *VIAF*, o gênero é apresentado de forma textual, por exemplo, *female* ou *male*, vinculado aos registros agregados de diferentes instituições. No *Wikidata*, essa informação é registrada por meio da propriedade P21 (*gender or sex*), que se associa a uma URI específica, como *wd:Q6581072* para *female* e *wd:Q6581097* para *male*, possibilitando a interoperabilidade e reutilização dos metadados em ambientes semânticos.

De acordo com Lira (2014), para efetuar o enriquecimento semântico e adquirir conceitos extras, utiliza-se uma variedade de recursos e técnicas, entre esses existe a anotação semântica que adiciona significado aos recursos, de forma automatizada ou manual; a vinculação e o mapeamento de recursos localiza links entre combinações semânticas dos dados e metadados a outros recursos *Web*; e a conversão para modelos de dados semânticos utilizando, por exemplo o *RDF/XML*.

Silva (2014) aponta que existem quatro tipos de anotações semânticas, a partir da utilização de *tags*, atributos, relações ou por ontologias. Dentre essas formas, no contexto da *Web* semântica, Hitzler (2021, p. 79) destaca que “[...] as ontologias são um dos principais meios para integração, compartilhamento e descoberta de dados, sendo uma ideia central que as próprias ontologias devem ser reutilizáveis por outros”, já que as ontologias são representações formais de domínios do conhecimento.

A falta de metadados detalhados e contextualizados é um desafio recorrente em diversas áreas do conhecimento, impactando diretamente a forma como as informações são representadas e recuperadas. A ausência de semântica pode dificultar a correlação de eventos históricos, tendências de escritas baseadas no gênero do autor, a análise de interações sociais ao longo do tempo e a descoberta de obras relacionadas a um mesmo contexto geográfico ou temporal. O enriquecimento semântico é um processo que pode auxiliar a resolver desafios, ao agregar significado e permitir a descoberta de informações de maneira mais eficiente e integrada.

3.2 Enriquecimento semântico de metadados LAM

O Quadro 8 sintetiza os 19 artigos que tratam dessa temática, organizados por título, autores, ano de publicação e categoria. A categorização foi realizada com base no foco principal de cada estudo, agrupando-os em áreas como Arquivos e Preservação de documentos, Bibliotecas Digitais e Repositórios e Museus e Patrimônio Cultural.

Quadro 8 - Artigos que abordam o enriquecimento semântico de metadados em LAM

Título	Autores	Categoria
Semantic Enrichment of Linked Archival Materials.	Chen, Shu-Jiun (2019)	Arquivos e Preservação de documentos
Impresso Inspect and Compare. Visual Comparison of Semantically Enriched Historical Newspaper Articles.	Düring, Marten <i>et al.</i> (2021)	
Impresso Text Reuse at Scale. An interface for the exploration of text reuse data in semantically enriched historical newspapers	Düring, Marten <i>et al.</i> (2023)	
Semantic Enrichment of Scientific Publications and Metadata.	Bertin, Marc; Atanassova, Iana (2012)	Bibliotecas Digitais e Repositórios
Automatic metadata generation in an archaeological digital library: Semantic annotation of grey literature	Vlachidi, Andreas <i>et al.</i> (2013)	
Semantic Enrichment: a Low-barrier Infrastructure and Proposal for Alignment.	Van Veen, Theo; Lonij, Juliette; Koppelaar, Hanna (2015)	
Capisco: low-cost concept-based access to digital libraries.	Hinze, Annika <i>et al.</i> (2019)	
Enriching metadata for a university repository by	Bellotto, Anna;	

modelling and infrastructure: a new vocabulary server for phaidra.	Bekesi, Janos (2019)	
Disponibilização de metadados em linked data para repositórios digitais.	Arakaki, Felipe <i>et al.</i> (2019)	
Automatic metadata extraction via image processing using Migne s Patrologia Graeca	Varthis, Evangelos <i>et al.</i> (2020)	
Semantic enrichment on large scanned collections through their "satellite texts": the paradigm of Migne's Patrologia Graeca	Varthis, Evangelos <i>et al.</i> (2022)	
A Linked Open Data model for describing comic book sequences: Exploring semantic enrichment opportunities with graphic medicine.	Petiya, Sean (2023)	
A recommendation system to promote local cultural heritage	di Bitonto, Pierpaolo <i>et al.</i> (2011)	Museus e Patrimônio Cultural
A systemic approach for effective semantic access to cultural content	Kollia, Ilianna <i>et al.</i> (2012)	
A linked open data framework to enhance the discoverability and impact of culture heritage.	Candela, Gustavo <i>et al.</i> (2019)	
3D point clouds in archaeology: Advances in acquisition, processing and knowledge integration applied to quasi-planar objects	Poux, Florent <i>et al.</i> (2017)	
Using type and temporal semantic enrichment to boost content discoverability and multilingualism in the Greek cultural aggregator SearchCulture.gr	Georgiadi, Haris <i>et al.</i> (2018)	
Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article.	Lei Zeng, Marcia (2019)	
Data modeling and evaluation of deep semantic annotation for cultural heritage images.	Wang, Xiaoguang <i>et al.</i> (2021)	

Fonte: elaborado pela autora (2025).

Os artigos selecionados discutem o enriquecimento semântico de metadados em Bibliotecas, Arquivos e Museus (LAMs). Diversas abordagens e métodos são empregados, incluindo o uso de ontologias, LOD, anotação semântica automatizada, fontes externas e *frameworks* de interoperabilidade.

Na primeira categoria, Arquivos Digitais e Preservação de documentos, Chen (2019) utiliza o fundo "*Chen Cheng-po's Paintings and Documents*" (CCP), pertencente aos Arquivos do Instituto de História de Taiwan para propor um modelo semântico para a conversão dos metadados em LOD. Para o desenvolvimento da ontologia de acordo com particularidades do domínio, adaptou o *Europeana Data Model* (EDM) e adotou quatro abordagens para enriquecimento semântico, o uso de vocabulários externos, reconciliação de dados, introdução de classes contextuais e extração de entidades nomeadas (NEE). Os resultados indicaram que a conversão

dos metadados arquivísticos para LOD melhora a acessibilidade, descoberta e reutilização das informações, além de facilitar sua integração a sistemas digitais.

Düring *et al.* (2021) apresentam a ferramenta *Inspect and Compare* (I&C), desenvolvida no projeto *Impresso*, voltada à exploração de jornais históricos digitalizados com apoio de técnicas de enriquecimento semântico, como reconhecimento e vinculação de entidades, modelagem de tópicos, classificação de artigos e detecção de reutilização textual. A ferramenta aprimora buscas, explorações de conteúdo e avaliação da qualidade dos dados, sendo bem avaliada apesar da interface complexa.

No estudo posterior, Düring *et al.* (2023) introduzem a interface *Text Reuse at Scale*, também vinculada ao projeto *Impresso*, que combina dados de reutilização textual com enriquecimento semântico. Utiliza o software *Passim* para detectar *clusters* de textos reutilizados, integrando modelagem de tópicos, reconhecimento de entidades e visualizações interativas. O estudo destaca a importância da detecção de reutilização de texto e do enriquecimento semântico para analisar a circulação de conteúdos históricos.

Na categoria de Bibliotecas Digitais e Repositórios, Bertin e Atanassova (2012) propõem o uso de anotações semânticas e extração automática para o enriquecimento de metadados em publicações científicas. O sistema desenvolvido aplica técnicas de processamento de linguagem natural (PLN) para analisar o contexto das citações, categorizando-as com base em uma ontologia linguística. Os resultados indicam que essa abordagem melhora a qualidade dos metadados e pode subsidiar o desenvolvimento de ferramentas para análise da produção científica. Os autores também ressaltam o crescimento do acesso aberto e a necessidade de aprimorar a análise de citações para promover um mapeamento mais preciso da ciência.

Na sequência, Vlachidis *et al.* (2013) apresentam um método para geração automática de metadados semânticos em literatura cinzenta arqueológica do *Archaeological Data Service* (OASIS), integrando o *Conceptual Reference Model* (CIDOC CRM), *CRM Extension for Archaeological Excavation* (CRM-EH) e PLN. Faz parte do projeto STAR, que busca melhorar a busca semântica e a interoperabilidade entre relatórios de escavação e bases de dados. A metodologia envolve pré-processamento, extração de informações e conversão para RDF, garantindo compatibilidade com padrões semânticos. Como resultado, a eficácia do método foi avaliada por especialistas, que validaram a precisão e a recuperação das

informações, demonstrando que a abordagem melhora significativamente o acesso e a integração de dados arqueológicos.

Van Veen, Lonij e Koppelaar (2015), descrevem a infraestrutura de enriquecimento semântico da Biblioteca Nacional dos Países Baixos (*KB Research*), que vincula entidades nomeadas de jornais históricos a bases externas. O reconhecimento ocorre de forma automática e manual, permitindo a edição e criação de enriquecimentos. O artigo propõe avanços como o uso de aprendizado de máquina, a substituição de *Freebase* por *Wikidata* e um ponto de extremidade SPARQL para ampliar a interoperabilidade. Concluem que a padronização de formatos, APIs e identificadores facilita o compartilhamento entre instituições, promovendo recuperação aprimorada da informação, integração eficiente e acesso simplificado a dados enriquecidos.

Hinze *et al.* (2019) apresentam o *Capisco*, um sistema de recuperação de informações de baixo custo que utiliza uma rede *Concept-in-Context* (CiC), gerada automaticamente a partir da *Wikipedia*, para enriquecimento semântico e indexação de documentos em bibliotecas digitais. A abordagem se baseia na associação de documentos a conceitos extraídos do contexto textual e na expansão semântica das consultas, permitindo desambiguação e inclusão de sinônimos e termos correlatos. Essa estratégia dispensa alterações estruturais nos sistemas, melhora a precisão da recuperação, reduz ruídos e torna os resultados mais alinhados ao significado buscado pelo usuário.

Bellotto e Bekesi (2019) descrevem a aplicação de enriquecimento semântico no repositório *Phaidra* da Universidade de Viena, com base em vocabulários controlados estruturados em RDF e conectados a recursos de *Linked Open Data* (LOD). A adoção da ferramenta iQvoc em substituição ao SKOSMOS resultou na padronização dos metadados e no controle de atributos como tipo de recurso, material e gênero dos recursos. A integração com vocabulários externos e ferramentas de consulta estruturada ampliou a interoperabilidade do repositório. Para manter a qualidade dos dados, os autores indicam a necessidade de diretrizes para manutenção, edição e expansão do vocabulário, além de documentação de apoio aos usuários.

Arakaki. F, *et al.* (2019) mostraram a viabilidade da conversão de dados de repositórios digitais para LD, com o objetivo de aprimorar a interoperabilidade e o enriquecimento semântico dos metadados. Foram identificadas bases de dados,

vocabulários e padrões a serem adotados para estruturar os dados conforme os princípios da *Web Semântica*. Como resultado, apresentaram um modelo de conversão que detalha as etapas de extração, transformação para RDF e enriquecimento semântico, utilizando ferramentas que possibilitam a integração com outros conjuntos de dados.

Varthis *et al.* (2020) propõem um método de extração automática de metadados semânticos em documentos digitalizados, utilizando *word spotting* como alternativa ao OCR tradicional. Aplicado à *Patrologia Graeca*, o método supera desafios como degradação de páginas, uso do grego politônico e segmentação de caracteres, alcançando 94,5% de precisão. Já Varthis *et al.* (2022) apresentam uma abordagem baseada no uso de textos satélites como índices e sumários para o enriquecimento semântico de coleções digitalizadas. A proposta resulta em um portal semântico estático, de baixo custo e escalável, com acesso facilitado por URIs intuitivas, otimizando a busca e recuperação da informação sem recorrer a mecanismos complexos de *Linked Data*.

No último artigo da categoria, Petiya (2023), introduz como o uso do LOD pode melhorar a descrição e a descoberta de histórias em quadrinhos (HQs), especialmente no campo da *graphic medicine* (medicina gráfica). Esse gênero inclui quadrinhos que abordam temas médicos e de saúde, ilustrando experiências de pacientes, doenças e tratamentos de maneira acessível e empática. No entanto, os metadados dessas obras frequentemente não incluem vocabulários médicos estruturados. O autor propõe um modelo baseado na *Comic Book Ontology* para mapear elementos dos quadrinhos (como páginas, quadros e balões) e conectá-los a ontologias médicas. Essa abordagem permite vincular partes específicas das HQs a conceitos médicos, melhorando sua indexação, descoberta e reutilização, especialmente para a educação em saúde e pesquisas sobre representação visual de doenças.

Na categoria Museus e Patrimônio Cultural, o primeiro artigo, Di Bitonto *et al.* (2011), descrevem um sistema de recomendação para o promover o turismo cultural do patrimônio cultural tangível e intangível. O objetivo do sistema é oferecer sugestões baseadas no perfil do usuário, enriquecidas semanticamente, para fornecer uma compreensão mais aprofundada do patrimônio cultural local. Os autores discutem as limitações dos sistemas existentes, que frequentemente se restringem a oferecer apenas informações básicas ou recomendam apenas um tipo de item. Já o sistema

proposto visa proporcionar uma experiência mais completa, incluindo materiais de aprofundamento para cada item recomendado.

Kollia *et al.* (2012) propõem um sistema voltado ao mapeamento de metadados culturais com foco na interoperabilidade e no enriquecimento semântico, por meio do alinhamento a modelos padronizados como o *Europeana Data Model* (EDM). A solução permite que instituições culturais integrem seus acervos com maior precisão semântica, utilizando técnicas de mapeamento automático de esquemas, representação ontológica em RDF e OWL, e execução de consultas por reescrita e inferência lógica. Testado com acervos gregos conectados à *Europeana*, o sistema demonstrou ganhos na qualidade e na abrangência da recuperação da informação.

Já Candela *et al.* (2019) apresentam a aplicação de técnicas de enriquecimento semântico na Biblioteca Virtual Miguel de Cervantes, com foco na padronização e interligação dos metadados. O processo incluiu a normalização de dados, a extração e desambiguação de entidades, e a vinculação a repositórios externos, promovendo maior interoperabilidade e enriquecimento contextual. O estudo, no entanto, evidencia desafios relevantes, como a heterogeneidade dos dados, variações linguísticas e limitações na qualidade das fontes externas, fatores que impactam a eficácia do processo de enriquecimento.

Poux *et al.* (2017) analisam os avanços da arqueologia digital e propõem a integração de nuvens de pontos 3D com enriquecimento semântico como estratégia para ampliar a análise e a preservação do patrimônio cultural. Dado que nuvens de pontos, por si só, carecem de significado semântico, os autores adotam uma metodologia fundamentada em ontologias formais, como o CIDOC-CRM, permitindo a classificação automática e a inferência de informações. Aplicado ao mosaico do oratório de *Germigny-des-Prés*, o método viabilizou a extração de dados e o desenvolvimento de uma plataforma web interativa. Os resultados apontam que a fusão entre representação geométrica e semântica melhora significativamente a recuperação e a interoperabilidade de dados arqueológicos.

Georgiadi *et al.* (2018) desenvolvem um esquema de enriquecimento semântico e normalização de metadados voltado ao SearchCulture.gr, agregador de conteúdos culturais do *National Documentation Centre* da Grécia. Diante da heterogeneidade dos metadados, os autores implementam uma abordagem baseada na plataforma Semantics.gr, utilizando vocabulários RDF para categorizar itens culturais e períodos históricos. As ações incluem normalização de datas, mapeamento

hierárquico de categorias e indexação semântica com filtros facetados e suporte à navegação interativa. Os dados foram convertidos para *Linked Data*, assegurando interoperabilidade com a *Europeana* e outros agregadores. Como resultado, observaram-se ganhos expressivos na descoberta e acessibilidade dos acervos, com suporte a linha do tempo interativa e navegação multilíngue.

Zeng (2019) descreve métodos para o enriquecimento semântico para melhorar a interoperabilidade, recuperação e reutilização dos dados LAM. A autora discute a necessidade de integrar tecnologias semânticas, como o LOD, ontologias e extração de entidades, para conectar e contextualizar informações de forma mais estruturada. São abordadas diferentes abordagens, incluindo a vinculação de metadados a bases externas, extração automática de entidades e uso de modelos ontológicos, permitindo consultas mais avançadas e acessibilidade aprimorada. Destaca ainda iniciativas como o *Europeana Semantic Enrichment Framework*, *Swissbib*, *MoMA*, *Digital Public Library of America (DPLA)* e *WorldCat Identities* que aplicam técnicas de enriquecimento semântico para melhorar a descoberta de informações e a interoperabilidade entre repositórios.

Por fim, Wang *et al.* (2021), propõem um modelo de anotação semântica profunda (DSA) para imagens de patrimônio cultural, uma abordagem que organiza informações em múltiplos níveis, conectando elementos visuais a conceitos semânticos e narrativas iconográficas. Baseado no método de *Panofsky*, a DSA interpreta as imagens em três níveis, aplicado aos murais das Cavernas de Mogao, utilizou ontologias como CIDOC-CRM e vocabulários controlados para garantir padronização, além de técnicas computacionais para extração automática de entidades e modelagem de relações semânticas. A DSA permite transformar imagens em LOD, facilitando sua recuperação e análise por meio de interfaces interativas. Os resultados demonstram que essa abordagem aprimora a pesquisa e acessibilidade de acervos culturais, tornando as imagens mais estruturadas e reutilizáveis nas Humanidades Digitais.

Como aponta Isaac *et al.* (2015), há uma diversidade de abordagens para o enriquecimento semântico, tornando inviável listar todas. No entanto, a partir dos 19 artigos analisados e apresentados no Quadro 8, foram sistematizadas no Quadro 9 as principais tecnologias semânticas identificadas na literatura, com o objetivo de explicitar os recursos empregados nos processos de enriquecimento semântico.

Quadro 9 - Tecnologias Semânticas

Tecnologias semânticas
Ontologias
Reconhecimento de Entidades Nomeadas (NER)
Linked Data
Vocabulários Controlados
Formalização em RDF/OWL
Indexação Semântica
Anotação Semântica
Modelagem de dados
Fontes externas
Inferência lógica
Processamento de linguagem natural (PNL)
Extração automática

Fonte: elaborado pela autora (2025).

A partir do Quadro 9, observa-se que o enriquecimento semântico, conforme evidenciado na literatura analisada, mobiliza um conjunto diversificado de tecnologias e metodologias que atuam de maneira complementar. De modo geral, os métodos apresentados demonstram a variedade de ferramentas aplicadas em diferentes domínios, refletindo a dinamicidade desse campo, que permanece em constante evolução. Essa diversidade responde à necessidade de estabelecer relações semânticas que qualifiquem os processos de representação, organização e recuperação da informação, promovendo maior interoperabilidade, contextualização e potencial de reutilização dos dados.

Os estudos analisados demonstram que a automatização do enriquecimento semântico tem permitido a ampliação da utilização dos dados, tornando o patrimônio cultural mais acessível. No entanto, os desafios ainda são significativos, principalmente em relação à heterogeneidade dos dados, variações linguísticas, granularidade dos metadados e a necessidade de padronização de vocabulários.

Outro aspecto é a necessidade de curadoria humana no processo de enriquecimento semântico. Embora métodos automatizados baseados em PLN, redes neurais e aprendizado de máquina tenham demonstrado eficiência na extração e vinculação de entidades, a supervisão de especialistas ainda é necessária para garantir a precisão das anotações, especialmente em documentos históricos e culturais, onde a interpretação semântica pode ser complexa, o que reforça a importância do profissional da informação.

O enriquecimento semântico vai além de um processo técnico dos sistemas de informação. Trata-se de um processo que visa agregar novos significados, estabelecer conexões e qualificar a representação de entidades por meio da integração de dados provenientes de múltiplas fontes, especialmente aquelas disponibilizadas na *Web* conforme os princípios do *Linked Data*.

O uso das ferramentas e estratégias bem estruturadas promove um acesso mais contextualizado e ligado ao conhecimento, fortalecendo o papel dos LAMs na preservação e disseminação da memória cultural e científica. O enriquecimento semântico, ao ampliar o contexto dos registros, permite que atributos, relacionamentos e contextos diversos sejam incorporados de forma mais dinâmica, precisa e interoperável.

Quando aplicado aos registros de autoridade, em especial, na representação do gênero, o enriquecimento semântico possibilita a vinculação desses registros a fontes externas confiáveis, como *Wikidata*, *VIAF* ou outras bases que mantêm informações atualizadas sobre identidade de gênero, o que amplia significativamente o alcance e a qualidade da representação.

A inclusão do metadado de gênero, deixa de ser uma prática isolada para se transformar em um componente integrado de uma rede informacional mais ampla, onde dados sobre identidade, atuação profissional, relações e outras características das entidades descritas podem ser associadas automaticamente, seguindo padrões abertos e interoperáveis. O uso de URIs garante a persistência e a desambiguação dessas informações, enquanto as ontologias e vocabulários controlados proporcionam a padronização semântica necessária para que sistemas distintos compreendam e processem os dados de forma uniforme.

Em especial, em relação a questão da representação da identidade de gênero, pois permite lidar com a sua natureza fluida, contextual e socialmente construída. A partir da *Web Semântica* e o enriquecimento semântico existe a possibilidade de uma representação contextualizada e atualizável da informação, assegurando maior respeito à diversidade identitária e às transformações que podem ocorrer ao longo da vida das pessoas descritas nos registros.

Além disso, o enriquecimento semântico favorece a interoperabilidade entre diferentes catálogos e sistemas de informação, garantindo que as informações sobre gênero, quando registradas de maneira qualificada, possam ser recuperadas,

cruzadas e analisadas em múltiplos contextos, ampliando as possibilidades de pesquisa, visibilidade e recuperação da produção intelectual.

Portanto, o enriquecimento semântico atua como um elemento mediador entre a necessidade técnica de interoperabilidade e a demanda social por representações mais inclusivas. Ele propicia uma infraestrutura informacional capaz de suportar a incorporação do gênero como metadado não de forma isolada ou estanque, mas como parte de um ecossistema informacional conectado, orientado por padrões abertos e tecnologias interoperáveis.

4. RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados obtidos a partir da análise dos registros de autoridade disponibilizados por bibliotecas nacionais, com foco na presença, estruturação e uso do metadado de gênero. Foram examinados catálogos que adotam diferentes padrões de metadados (MARC 21, UNIMARC e INTERMARC), a fim de identificar como as instituições têm tratado a representação da identidade de gênero em seus registros. Os dados analisados permitiram mapear os campos utilizados, os valores atribuídos, os vocabulários empregados e a eventual vinculação a identificadores externos, como VIAF e *Wikidata*.

Além disso, buscou-se identificar diretrizes institucionais, políticas públicas e práticas locais que fundamentam essas escolhas, bem como destacar as boas práticas observadas ao longo do levantamento. Os resultados são discutidos à luz da literatura especializada e dos referenciais técnicos apresentados nos capítulos anteriores, articulando aspectos descritivos, éticos e tecnológicos da inclusão do metadado de gênero nos sistemas de informação.

4.1 Levantamento sobre à utilização do metadado de gênero em registros de autoridade

A análise qualitativa dos registros de autoridade nos catálogos das bibliotecas nacionais evidenciou padrões recorrentes que indicam uma aplicação limitada e desigual do metadado de gênero. Após a coleta dos dados e o contato com as instituições, foram incluídas no Quadro 10 apenas as bibliotecas nacionais que responderam ao e-mail enviado e/ou aquelas cujos registros permitiram verificar a presença do campo referente ao gênero.

Em primeiro lugar, destaca-se a prevalência do modelo binário na maioria dos registros que apresentam essa informação. As designações “feminino” e “masculino” aparecem como as únicas opções utilizadas, mesmo em contextos em que o formato MARC 21 permite o uso de vocabulários controlados mais abrangentes, fontes externas e subcampos voltados à explicitação da diversidade identitária. Essa padronização restritiva demonstra a ausência de um olhar ampliado sobre a questão

do gênero enquanto construção sociocultural e relacional, refletindo práticas descritivas ainda centradas em categorias fixas.

Além da limitação dos valores empregados, observou-se também uma significativa lacuna quanto ao preenchimento do campo 375. Em muitas bibliotecas, o campo encontra-se ausente mesmo em registros que incluem informações como nacionalidade, ocupação, campo de atividade e idioma. Essa ausência sugere que, embora existam condições técnicas para a inclusão do metadado de gênero, sua aplicação ainda não está incorporada às rotinas de catalogação, seja por falta de diretrizes institucionais claras, seja por resistências no plano das políticas descritivas. Essa omissão compromete a representatividade e a possibilidade de recuperação informacional por atributos identitários, afetando, sobretudo, autoras e sujeitos cuja identidade de gênero foge às categorias hegemônicas.

Por outro lado, em menor número, algumas instituições vêm adotando práticas que demonstram maior alinhamento com os princípios da *Web Semântica*. Entre essas boas práticas, destaca-se o uso de URIs vinculadas a fontes confiáveis como VIAF e *Wikidata*, bem como a presença dos subcampos \$u (identificador externo), \$v (fonte da informação), \$2 (código do vocabulário utilizado) e \$1 (URI do objeto do mundo real). Esses elementos indicam um esforço por parte dessas bibliotecas em qualificar a representação do gênero com base em fontes externas, aumentando a granularidade, a interoperabilidade e a rastreabilidade dos dados. Ainda assim, esses casos permanecem pontuais e carecem de uniformidade em sua aplicação, indicando que a adoção de abordagens semânticas ainda se encontra em estágio incipiente nos registros de autoridade analisados.

Essa heterogeneidade reforça a necessidade de diretrizes integradas que articulem os recursos técnicos disponíveis nos formatos de metadados com compromissos sociais orientados à inclusão e à pluralidade identitária. Sendo o Quadro 10, representa um levantamento de bibliotecas nacionais quanto à utilização do metadado de gênero em registros de autoridade.

É importante destacar que foram observados aspectos como: o campo e indicador utilizado, os valores adotados para representar o gênero, a decisão institucional quanto ao registro dessa informação.

Quadro 10 - Bibliotecas Nacionais que utilizam ou não o metadado de gênero
Américas

Biblioteca	Campo e subcampo utilizado	Valores	Decisão sobre o registro do gênero
Biblioteca A	Campo 375 subcampo \$a do MARC 21	Mulher Homem Mulher trans Homem trans Intersexo Travesti Não binário Outro Não especificado	Registra o gênero e ampliou os valores.
Biblioteca B	Não utiliza - informação confirmada via catálogo e contato institucional	-	O registro de autoridades ainda não foi implementado.
Biblioteca C	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero. Para registro de nomes pessoais utiliza: 100, 400, 500, 670 e 675.
Biblioteca D	Não utiliza - informação confirmada via catálogo e contato institucional	-	Utilizou o campo 375 entre 2018 e 2019, mas interrompeu o uso por questões de privacidade. Com a recomendação do PCC em 2022, considerou sua prática já alinhada à orientação de não registrar o gênero, embora alguns registros antigos ainda contenham a informação.
Biblioteca E	Não utiliza - informação confirmada via catálogo e contato institucional	-	Utilizou o campo 375, mas como membro do PCC segue suas diretrizes. Desde 2022, adota a política de não registrar o campo e de removê-lo durante a edição de registros.
Biblioteca F	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero
Biblioteca G	Campo 375 subcampo \$a do MARC 21	M (masculino) F (feminino) U (desconhecido) X (não aplicável)	Implementou, em 2021, o uso do campo 375 do MARC 21 para registrar o gênero dos autores nos registros de autoridade, seguindo as diretrizes estabelecidas no RDA.
Biblioteca H	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	O gênero é registrado como feminino ou masculino, não sendo utilizada a opção "desconhecido". Para casos em que o autor se identifica com outro gênero, registra-se um termo não controlado ou transcrito da fonte consultada.
Biblioteca I	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo 375 (Gênero), mas em breve, seguindo as normativas do RDA, passará a utilizá-lo, motivo pelo qual se encontra analisando as diretrizes específicas para seu uso adotadas por outras Bibliotecas Nacionais.

Biblioteca J	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	Segue as diretrizes do <i>Manual de control de autoridades de nombre de personas aplicando RDA</i> . Embora haja a identificação, recuperação e geração de relatórios por gênero, uma política interna determina que a informação não seja exibida no catálogo de autoridades.
Biblioteca K	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	Utiliza apenas os códigos da ISO 5218: "m" (masculino) e "f" (feminino), sem uma opção específica para gênero diverso, pois os códigos "0" e "9" representam "desconhecido" ou "não aplicável" e não são considerados adequados.
Biblioteca L	Campo 375 subcampo \$a do MARC 21	M (masculino) F (feminino) U (desconhecido) X (não aplicável).	Segue as diretrizes estabelecidas no RDA.
Biblioteca M	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero.
Biblioteca N	Não é possível verificar. informação confirmada via catálogo e contato institucional	-	Sofreu um ataque cibernético em 2023, comprometendo sua estrutura e serviços. Atualmente, não há acesso ao sistema de catalogação, impossibilitando a verificação de informações, e não há fontes alternativas disponíveis.
Biblioteca O	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	Segue as diretrizes estabelecidas no RDA. No CANTIC, esse campo é opcional, ficando a critério de cada instituição participante. Não há uma lista normalizada de identidades de gênero. Caso necessário, pode-se usar outro termo de um vocabulário controlado.
Biblioteca P	Não utiliza - informação confirmada via catálogo e contato institucional	-	O sistema ALMA vincula os campos 100 e 700 ao LC NAMES. Se o campo 375 estiver presente no LC NAMES, os registros bibliográficos referenciam essa informação, mas a biblioteca não a insere ativamente.
Biblioteca Q	Não utiliza - informação confirmada via catálogo e contato institucional	-	O campo 375 foi utilizado por um curto período, mas deixou de ser registrado desde a diretriz do PCC de abril de 2022. Atualmente, a biblioteca não inclui o campo 375 em novos registros e remove esse campo ao atualizar registros existentes.
Biblioteca R	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	Utiliza o campo 375 com os valores "Masculino" e "Femenino" (não são contemplados outros valores).
Biblioteca S	Campo 120 subcampo \$a do CONMARC/A	"a" feminino "b" masculino "c" transgênero	No sistema COBISS, o formato COMARC/A é utilizado para registros de autoridade, baseado no

		"u" desconhecido ou não determinável)	UNIMARC/A. O gênero da entidade é registrado no subcampo 120\$a.
Biblioteca T	Não utiliza - informação confirmada via catálogo e contato institucional	-	Apenas registram os nomes das pessoas (e variações dos nomes, se houver) e a data de nascimento.
Biblioteca U	Não utiliza - informação confirmada via catálogo e contato institucional	-	Utilizou o campo 375 e desde 2022, adota a política de não registrar o campo e de removê-lo durante a edição de registros.
Biblioteca V	INTERMARC Campo 17 008 UNIMARC 120\$a	F (feminino) M (masculino) # (não informado) a = feminino B = masculino	Os dados são produzidos em INTERMARC e também divulgados em UNIMARC. No INTERMARC, o gênero é registrado na posição 17 do campo 008, com os valores: f (feminino), m (masculino) e # (não informado). No UNIMARC, a informação aparece no campo 120 \$a, com a para feminino e b para masculino.
Biblioteca W	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero.
Biblioteca X	Não utiliza - informação confirmada via catálogo e contato institucional	-	Utiliza o DanMarc2 e o campo 375 não existe.
Biblioteca Y	Não utiliza - informação confirmada via catálogo e contato institucional	-	Utiliza o HUNMARC e o campo 375 não existe. Está em transição para o MARC21, no qual o campo será adotado.
Biblioteca Z	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não cria registros detalhados de autoridade. Contém dados básicos como o nome conforme aparece na obra, com campos 670 e 675, além de datas de nascimento e/ou morte, se disponíveis.
Biblioteca AA	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino) Outros	Os valores usados são masculino, feminino e outros, se houver alguma dúvida sobre qual valor usar, o campo não será incluído no registro de autoridade.
Biblioteca AB	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero.
Biblioteca AC	Não utiliza - informação confirmada via catálogo e contato institucional	-	O uso de registros de autoridade é limitado ao nome e, quando disponível, datas biográficas e títulos. É mantido no mínimo, com a intenção de identificar uma pessoa, em vez de criar um perfil completo da pessoa. Até agora, esses dados

			foram suficientes para identificar um autor de outro. Caso precise adicionar mais informações para identificar uma pessoa de outra, pode-se utilizar outros campos, como no caso do campo 375.
Biblioteca AD	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	Utiliza o formato MARC e inclui o campo 375 nos registros de autoridade. Os valores aceitos para o subcampo \$a são "f" (feminino) e "m" (masculino). O campo é repetível em casos de mudança de gênero, e podem ser usados os subcampos \$s (data inicial) e \$t (data final) para indicar o período de identificação com determinado gênero.
Biblioteca AE	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero.
Biblioteca AF	Não utiliza - informação confirmada via catálogo e contato institucional	-	Não utiliza o campo para o registro de gênero.
Biblioteca AG	Campo 375 subcampo \$a do MARC 21	F (feminino) M (masculino)	Utiliza o campo 375 com os valores "Masculino" e "Feminino" (não são contemplados outros valores).
Biblioteca AH	Não utiliza - informação confirmada via catálogo e contato institucional	-	Registravam o gênero, porém, foram removidas as informações de gênero dos registros de autoridade por serem membros do ISNI e não permitirem informações de gênero.

Fonte: Elaborado pela autora (2025).

A relação entre os padrões de metadados, esquemas de descrição, as decisões tomadas para a inclusão do metadados de gênero nos registros de autoridade e os valores registrados possibilitam compreender as variações entre as bibliotecas nacionais analisadas. O RDA Oficial, enquanto diretrizes voltadas à descrição de recursos informacionais com base em princípios da *Web Semântica* e na modelagem orientada por entidades, oferece suporte ao registro do gênero no elemento 9.7. Esse elemento prevê a possibilidade de indicar o gênero com o qual a pessoa se identifica, com os valores M (masculino), F (feminino), U (desconhecido) e X (não aplicável).

As bibliotecas A, G, H, J, K, L, O, R, AA, AD e AG utilizam o campo 375 do MARC 21, criado especificamente para registrar o atributo de gênero em registros de autoridade. No entanto, apesar de utilizarem o mesmo campo, essas instituições

optaram por valores diferentes. Por exemplo, a Biblioteca A, opta por uma lista expandida de termos que contempla identidades de gênero diversas, como “Mulher”, “Homem”, “Mulher trans”, “Homem trans”, “Intersexo”, “Travesti”, “Não binário” e “Outro”.

Já as bibliotecas G, J, L e AA utilizam os valores M (masculino), F (feminino), U (desconhecido) e X (não aplicável), conforme previsto no RDA. Por outro lado, entre as bibliotecas que também utilizam o campo 375, mas com aplicação mais restrita, destacam-se H, K, R, O, AD e AG, nas quais os valores são “masculino” e “feminino”, sem contemplar outras possibilidades de identificação de gênero.

Entre as instituições, a Biblioteca K utiliza o campo 375 com os códigos “M” (masculino) e “F” (feminino), baseando-se na ISO 5218. Essa biblioteca também explicita que os códigos “0” (desconhecido) e “9” (não aplicável) não são utilizados, demonstrando um controle sobre os valores empregados. Essas diferenças evidenciam que, embora o campo 375 esteja disponível e implementado, sua utilização está condicionada por decisões institucionais, políticas de catalogação locais e pela interpretação dos esquemas de descrição como o RDA, no contexto de cada biblioteca.

É importante destacar também as bibliotecas que adotam outros formatos além do MARC 21, como as bibliotecas S e V. A Biblioteca S utiliza o CONMARC/A, e registra o gênero no campo 120, subcampo \$a, com os seguintes valores: "a" (feminino), "b" (masculino), "c" (transgênero) e "u" (desconhecido ou não determinável). Já a Biblioteca V emprega uma estrutura combinada, operando com o INTERMARC e o UNIMARC. Nesse caso, o gênero é registrado no campo 17 008 do INTERMARC e no campo 120 \$a do UNIMARC, utilizando os valores: "F" (feminino), "M" (masculino), "#" (não informado), "a" (feminino) e "b" (masculino). Esses exemplos demonstram como diferentes formatos adotam campos diferentes para o metadado de gênero e valores.

Algumas bibliotecas já utilizaram o campo 375, mas optaram por descontinuar seu uso. É o caso das bibliotecas D, E, N, Q e U, que registravam o metadado de gênero em períodos anteriores, mas passaram a alinhar suas práticas à recomendação formal do *Program for Cooperative Cataloging* (PCC), emitida em 2022, que orienta as instituições a não utilizarem mais o campo e a removê-lo sempre que os registros forem editados. A Biblioteca AH utilizou o campo 375 em registros

anteriores, porém, como integrante da rede ISNI, passou a seguir as diretrizes, que orienta a não inclusão do gênero.

As bibliotecas B, C, F, I, M, P, T, W, X, Y, Z, AC, AE e AF não utilizam metadado de gênero e, até o momento, nunca o adotaram em seus registros de autoridade. Essa informação foi verificada por meio da análise direta dos catálogos e confirmada por contato institucional via e-mail. No caso da Biblioteca N, não foi possível confirmar a presença ou ausência do campo no catálogo, uma vez que a instituição sofreu um ataque cibernético.

Para ilustrar como esse metadado tem sido aplicado nos registros de autoridade, serão apresentados a seguir exemplos de registros de autoras em diferentes catálogos nacionais.

4.2 Iniciativas dos catálogos que incluem o metadado de gênero

Para ilustrar como esse metadado tem sido aplicado nos registros de autoridade, serão apresentados a seguir exemplos de registros das autoras Agatha Christie⁵, J. L. Butler⁶, Jane Austen⁷, Lola Anglada⁸ e Virginia Woolf⁹ em catálogos nacionais selecionados. A escolha dessas autoras considera sua ampla presença em diferentes bases de autoridade, além de sua representatividade em contextos históricos, culturais e linguísticos diversos.

A análise desses registros permite compreender como o metadado de gênero é estruturado e de que forma as bibliotecas nacionais têm incluído essa informação nos registros de autoridade.

⁵ Agatha Christie (1890–1976) é uma das autoras mais traduzidas e publicadas do século XX, amplamente reconhecida por suas obras de ficção policial.

⁶ J. L. Butler é um pseudônimo da autora Tasmine Perry cujo romances são best-sellers.

⁷ Jane Austen (1775–1817), autora inglesa do período georgiano, tem presença consolidada na literatura clássica ocidental.

⁸ Lola Anglada (1892–1984) foi escritora, ilustradora e ativista catalã, figura de destaque na literatura infantil e na produção cultural da Catalunha.

⁹ Virginia Woolf (1882–1941), autora britânica associada ao modernismo literário.

4.2.1 Biblioteca Nacional Mariano Moreno (BNMM)

A Biblioteca Nacional Mariano Moreno (BNMM)¹⁰, da Argentina, adota o formato MARC 21 de autoridade. A Figura 5 apresenta o registro da autora Jane Austen, no qual é possível observar a presença do campo 375, destinado ao registro do gênero.

Figura 5 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional Mariano Moreno

```

Cerrar formato MARC
FMT AU
LDR cz a22 n 4500
001 000021711
003 AR-BaBN
005 20221201132756.0
008 790418n| acannaabn |b aaa
0247 |a 102333412 |2 viaf |1 http://viaf.org/viaf/102333412
040 |a DLC |b eng |c DLC |d DLC |d C|Y-BR |d OCoLC |d DLC-R |d SG-SILA |d OCoLC |d DLC |d CSt |d UPB |d AR-BaBN
043 |a e-uk-en
046 |f 17751216 |g 18170718
1001 |a Austen, Jane, |d 1775-1817
370 |a Steventon (Hampshire, Inglaterra) |b Winchester (Inglaterra) |e Bath (Inglaterra)
374 |a Autores
375 |a Mujer ←
377 |a eng |l inglés
4001 |a Ostin, Dzhein, |d 1775-1817
4000 |a Ao-ssu-ting, |d 1775-1817
4000 |a Aosiding, |d 1775-1817
4001 |a Ao-ssu-ting, Chien, |d 1775-1817
4001 |a Aosiding, Jian, |d 1775-1817
4001 |a Austenová , Jane, |d 1775-1817
670 |a Wikipedia, via WWW, Feb. 8, 2012: |b (n. dic. 16, 1775 en Steventon, Hampshire; m. 18 jul. 1817 en Winchester, Hampshire; novelista inglesa; 1800 se traslada a Bath)
670 |a Persuasión, 1978.
6780 |a Nació en 1775 y falleció en 1817, en su hogar contaba con una biblioteca privada que la sumergió en la lectura, su talento como escritora aflora en su adolescencia, fué una destacada novelista británica. Algunas de sus obras son llevadas al cine. La mayoría de sus obras aparecen bajo seudónimo.
SYS 000021711

```

Fonte: Biblioteca Nacional da Argentina (2025).

Entre os diversos campos exibidos, destaca-se o campo 375, o subcampo \$a, que indica o gênero da entidade descrita com o valor "Mujer" (Mulher). O campo 100, apresenta o nome autorizado "Austen, Jane, 1775-1817", o campo 370, que indica locais associados à autora, como Steventon, Winchester e Bath, e o campo 374, que registra sua ocupação como "Autores". Além disso, o campo 377 informa a língua

¹⁰ Registro Jane Austen Biblioteca Nacional da Argentina:
https://catalogo.bn.gov.ar/F/P7SAD7ATR4EL9KCL4JIQK15GSRT2AE1HTJ2SQ5K14RVFJ2GUSE-15456?func=find-acc&acc_sequence=000391649.

associada, especificando o inglês, enquanto os campos 400 e 500 listam variações do nome da autora. O campo 670 traz referências externas, incluindo a *Wikipédia*, e o campo 678 contém uma nota biográfica que destaca a importância de Jane Austen como escritora.

Observa-se a aplicação de técnicas semânticas que contribuem para a interoperabilidade dos dados de autoridade. Um dos principais indicativos dessa aplicação é o uso do campo 0247, que contém um identificador VIAF vinculado à entidade descrita, acompanhado da URI <http://viaf.org/viaf/102333412>. Esse campo estabelece uma ligação semântica externa entre o registro local da biblioteca e uma base de autoridade internacional, permitindo que sistemas interoperáveis reconheçam e reutilizem a informação de forma automatizada.

Conforme o contato com BNMM, os registros de autoridade são criados de acordo com o AACR2 no que se refere à formulação e escolha do cabeçalho autorizado ou forma preferida. No entanto, considerando o avanço em direção ao RDA, em 2015 foram adotadas diretrizes para a utilização do campo 375¹¹. Desde então, optou por hibridar os registros de nomes pessoais com os campos de atributos previstos no modelo conceitual FRAD, entre eles os campos 370, 372, 373, 374, 375 e 377. Inicialmente, o campo 375 foi utilizado com os valores “masculino” e “feminino”, conforme previsto na norma ISO.

Durante a pandemia, a BNMM realizou uma campanha de levantamento de dados biográficos junto a autores, com o objetivo de obter informações de primeira mão. Nesse contexto, a Comissão de Gênero da instituição, com base nos lineamentos do Ministério das Mulheres da Argentina, propôs a ampliação dos termos utilizados no campo de gênero.

A BNMM compreende as justificativas apresentadas pelos comitês de trabalho do RDA Original, que consideram o campo 375 como pertencente à esfera da privacidade pessoal, razão pela qual recomendam sua exclusão dos atributos descritivos. No entanto, de acordo com a Lei de Proteção de Dados Pessoais, o uso dessas informações é permitido quando autorizado pela própria pessoa ou quando obtido a partir de fontes de acesso público. Diante disso, a Divisão de Controle de Autoridades, vinculada ao Departamento de Processos Técnicos, optou por manter a utilização do campo 375 nos registros de autoridade da instituição. Ainda a instituição

¹¹ Manual de procedimento BNMM: <https://escritoriopt.bn.gov.ar/mcat/mcata-3XX.html>.

defende que o uso desse dado contribui para conhecer a composição da produção bibliográfica por gênero, o que pode subsidiar novas linhas de pesquisa e análise na área.

4.2.2 Biblioteca Nacional da Colômbia

Na sequência, a Figura 6 apresenta o registro de autoridade da autora *Virginia Woolf* no catálogo da Biblioteca Nacional da Colômbia (BNC).

Figura 6 - Registro de autoridade de Virginia Woolf no catálogo da Biblioteca Nacional da Colômbia

The screenshot displays the authority record for Virginia Woolf. At the top, there is a blue header with the text 'Detalles de AutoridadAutoridad' and a close button. Below this, a search button labeled 'Buscar en catálogo' is visible. The record includes a profile icon, a 'Share' button, and a dropdown menu 'Seleccionar una acción'. The main details are as follows:

Tipo de Autoridad: Nombre personal	Nombre/Título: ✓
Nombre personal autorizado: Woolf, Virginia, 1882-1941	Serie: ✗
Lugar asociado: Kensington (Londres, Reino Unido)	Materia: ✓
Género: Femenino	Subdivisión: ✗

Fonte: Biblioteca Nacional da Colômbia (2025).

Nesse exemplo, observa-se a presença do metadado de gênero explicitamente indicado como “Femenino” (feminino), demonstrando que a instituição adota o campo de gênero como parte da descrição da entidade. O registro indica que se trata de um nome pessoal, apresentando o nome autorizado da autora como “Woolf, Virginia, 1882-1941” e informações associadas, como o lugar de nascimento, especificado como *Kensington*, Londres, Reino Unido. Uma particularidade desse registro é que não há a opção de visualizar os dados no formato MARC 21, impossibilitando a análise detalhada dos campos e subcampos utilizados para sua estruturação.

Na Figura 7, é possível filtrar por gênero¹², o que implica a existência de um campo estruturado para esse atributo nos registros de autoridade. Trata-se de uma estratégia que contribui para a construção de catálogos mais descritivos,

¹² Catálogo de autoridades BNC:
https://bnco.ent.sirsi.net/client/es_ES/autoridades/search/results?te=XREF.

transparentes e potencialmente alinhados a práticas inclusivas de representação, ao reconhecer o gênero um atributo relevante na descrição de pessoas.

Figura 7 - Registro de autoridade de Virginia Woolf no catálogo da Biblioteca Nacional da Colômbia

The screenshot displays the authority record interface for Virginia Woolf in the National Library of Colombia's catalog. On the left, there are filters for 'Lugar asociado' (Colombia: 7218, Colombia: 571, Estados Unidos: 554, España: 361, México: 253), 'Tipo de serie' (Otro tipo de serie: 1874, Serie monográfica: 491), and 'Género' (Masculino: 10672, Femenino: 8635, Masculino: 2432, Femenino: 1918, Maculino: 3). The main area shows four records (6, 7, 8, 9) with buttons for 'Mostrar detalle de autoridad' and 'Buscar en el catálogo'. Record 6: 'Nombre personal autorizado: Bautista Quintero, Pedro E.'; Record 7: 'Nombre personal autorizado: Amunátegui y Solar, Domingo, 1860-1946'; Record 8: 'Nombre personal autorizado: Bautista Urbaneja, Diego, 1782-1856'; Record 9: 'Nombre personal autorizado: Amurri, Antonio'. Each record has status indicators for 'Nombre/Título', 'Serie', 'Materia', and 'Subdivisión'.

Fonte: Biblioteca Nacional da Colômbia (2025).

Não se observam explicitamente URIs do VIAF e *Wikidata* nem elementos típicos da *Web Semântica*, como vocabulários controlados com referência externa, subcampo \$1, ou exposição em RDF.

4.2.3 Biblioteca Nacional do México

Na Figura 8, está o registro de autoridade da autora Jane Austen na Biblioteca Nacional do México (BNM).¹³

¹³ Catálogo de autoridade Biblioteca Nacional do México:

https://catalogo.iib.unam.mx/F/D4T3N22VPEGYLNYX8SSEJGVJAHXRIM7LG86DPMV1RQQHEXK7PN-13627?func=full-set-set&set_number=010962&set_entry=000001&format=001.

Figura 8 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional do México

The screenshot shows the Biblioteca Nacional website interface. At the top, there is a navigation bar with the logo and the text "BIBLIOTECA NACIONAL". Below this, there are links for "Búsquedas", "Índices", "Lista de resultados", "Ir a Catálogo Nautilo", and "Terminar". A secondary bar contains "Agregar a mi estante electrónico", "Guardar/Correo", and "Guardar en servidor".

The main content area is titled "Vista del registro" and shows the following MARC 21 record:

```

Seleccione formato: Estándar Nombres de etiquetas Etiquetas MARC
Registro 1 de 1
FMT AU
LDR nz a n 4500
001 000067837
005 20170810094631.0
008 140313 ||azannaabn|||||||n|aaa|||
040 |a MX-MxBN |b spa |e rda |c MX-MxBN
046 |f 1775 |g 1817
1001 |a Austen, Jane, |d 1775-1817
370 |a Steventon (Inglaterra) |b Winchester (Inglaterra)
372 |a Literatura |a Literatura inglesa |2 BNMx
374 |a escritor |2 toit
375 |a femenino ←
377 |a eng
670 |a BNM. Base bibliográfica, mar. 13, 2014: |b (Austen, Jane, 1775-1817)
670 |a Orgullo y prejuicio, 2013: |b portada (Jane Austen) verso de la portada (nació en Steventon, Inglaterra en 1775; murió en 1817)
670 |a Orgullo y prejuicio, 2016: |b portada (Austen, Jane) contraportada (Biografía... (Steventon, 1775 - Winchester, 1817) es la primera gran mujer de la literatura inglesa... punto de referencia ineludible en la historia de la literatura universal.) verso de la portada (Título original: Pride and prejudice)
70010 |a Austen, Jane, |d 1775-1817 |0 n 79032879
SYS 000067837
  
```

Fonte: Biblioteca Nacional do México (2025).

Dentre os campos exibidos, destaca-se o campo 375, que indica o gênero da autora com o valor "femenino" (feminino), demonstrando a adoção do metadado de gênero como parte da descrição da entidade.

Além desse campo, o registro contém campo 100 registra o nome autorizado como "Austen, Jane, 1775-1817", incluindo as datas de nascimento e falecimento. O campo 370 informa os locais associados à autora, indicando *Steventon* (Inglaterra) como local de nascimento e *Winchester* (Inglaterra) como local de falecimento. No campo 372, a atividade associada à entidade é registrada como "Literatura", enquanto o campo 374 especifica a ocupação de "escritora". Já o campo 377 indica a língua associada, neste caso, inglês. O registro também inclui o campo 670, com referências às obras da autora, como *Orgullo y Prejuicio* (Orgulho e Preconceito).

Em contato com a BNM, a instituição segue as regras do RDA Toolkit em sua versão oficial para o uso do campo 375, registrando o gênero da entidade que inclui as opções "femenino" e "masculino", optando por não utilizar "desconocido". No caso de autores que se identificam com outro gênero, a BNM permite o uso de termos não controlados ou a transcrição do termo conforme aparece no recurso consultado.

Embora o registro seja tecnicamente estruturado, não se observam no registro elementos característicos de técnicas semânticas no sentido estrito como o uso

de identificadores persistentes (URI) no campo 024, ou a vinculação com bases externas como VIAF e *Wikidata*. Todos os campos estão preenchidos com valores textuais, e não há subcampos que apontem para vocabulários controlados com codificação referenciável (ex.: subcampo \$1, \$0, ou \$2).

4.2.4 Biblioteca Nacional do Peru

Na Figura 9, está o registro de autoridade da autora Jane Austen no catálogo de autoridades da Biblioteca Nacional do Peru¹⁴ (BNP).

Figura 9 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional do Peru

005	20100319:10304500
008	100319 baabzaaaa a a aaa d
040	\$a pelibn \$b spa \$e rca \$c pelibn
046	\$f 1775 \$g 1817
100	1 \$a Austen, Jane, \$d 1775-1817
670	\$a Internet, http://www.bnf.fr , 19 dic. 2005 : \$b página de la Bib. Nac. de Francia (Austen, Jane, 1775-1817 ; n. Steventon, Gran Bretaña, 12 dic. 1775-m. Winchester, Gran Bretaña, 18 jul. 1817 ; novelista)

Formato Visualización MARC

<https://info.bnp.gob.pe/cgi-bin/autoridades/OHisjhx0MeI7N3HSPGEvSXNKN3/NT3?ACC=165&DOC=1#docu>

Fonte: Biblioteca Nacional do Peru (2025).

O registro contém elementos descritivos, como o campo 100, que apresenta o nome autorizado da autora acompanhado das datas de nascimento e morte, e o campo 670, que indica a fonte consultada para validação das informações. No entanto, o campo 375, destinado ao registro do metadado de gênero, não é exibido na interface pública do catálogo.

Em contato com a instituição, foi informado que o campo 375 é utilizado internamente nos registros de autoridade. Contudo, por decisão administrativa, a BNP

¹⁴ Catálogo de autoridades da BNP: <https://info.bnp.gob.pe/cgi-bin/autoridades/OTIvvhG6krOJ4TnXcF7BQ139J4/NT1?ACC=165&DOC=1>.

opta por não tornar essa informação visível ao público na interface do catálogo. Essa ação visa equilibrar práticas de descrição com considerações institucionais e éticas relativas à exposição de atributos pessoais. Apesar da ocultação, a presença do campo no sistema permite a realização de consultas internas, relatórios e análises baseadas em gênero, o que indica que o metadado continua ativo em nível técnico e funcional.

Na Figura 10, observa-se que a BNP aplica técnicas semânticas em seu catálogo de autoridades. A interface apresenta um recurso de vinculação automática, por meio do gerador de enlaces “AbsysNet c+/link”, que conecta o registro de autoridade a fontes externas. No exemplo analisado, o registro da autora Jane Austen está vinculado diretamente a identificadores persistentes em bases como VIAF, ISNI, *Wikidata* e a base de dados da Biblioteca Nacional da Espanha (Datos BNE). A existência desses vínculos demonstra a adoção de práticas associadas à *Web Semântica*, permitindo que os registros da BNP se integrem a redes internacionais de autoridade e contribuam para a interoperabilidade entre catálogos.

Figura 10 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Nacional do Peru

The screenshot displays the 'Catálogo de autoridades' interface for Jane Austen. At the top, there is a search bar and navigation links for 'Inicio', 'Marcados', and 'Ayuda'. Below the search bar, there are two buttons: 'Más información' and 'Quizás le interese:'. The main content area is divided into several sections:


- Búsqueda en el catálogo:** Includes a search bar and a link for 'Obras asociadas' (Todas las publicaciones en la biblioteca).
- Generador de enlaces AbsysNet c+/link:** A section for generating automatic links to internet resources, with the text 'Generando enlaces en Internet para'.
- Datos BNE:** A box showing the BNE logo and the text 'Datos BNE Austen, Jane,'.
- VIAF - Fichero de Autoridades Virtual Internacional:** A box showing the VIAF logo and the text 'VIAF Austen, Jane,'.
- ISNI - Identificador Estándar Internacional de Nombres:** A box showing the ISNI logo and the text 'ISNI Austen, Jane,'.
- Wikidata:** A box showing the Wikidata logo and the text 'Wikidata Austen, Jane,'.

Fonte: Biblioteca Nacional do Peru (2025).

4.2.5 Biblioteca Nacional da Alemanha

Na figura 11, é apresentado o registro de autoridade de *Jane Austen* no catálogo da Biblioteca Nacional da Alemanha (DNB)¹⁵, dentro da estrutura da *Gemeinsame Normdatei* (GND), o arquivo normativo comum utilizado por instituições alemãs e austríacas para controle de autoridade.

Figura 11 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Nacional Alemanha

GND	
Link zu diesem Datensatz	https://d-nb.info/gnd/118505173
Person	Austen, Jane
Geschlecht	weiblich 
Andere Namen	Aosiding (Chines.) Aosiding, Jian Ao-ssu-ting, Chien Ao-ssu-ting Austenová, Jane Austenové, Jane Austenová Jane Awsten, Jeyn (Uigur.) Osten, Dzhejn Остен, Жейн Ostin, Dzhejn (Russ. Namensform, AACR) Ostin, Džejn (Russ. Namensform, RAK-WB) Ostin, Zhejn Остин, Жэйн Ostjen, Zhjejn 奧斯丁 (Schriftcode: Hans) 奧斯婷 (Chinesische Namensform (Transkription um 1950)) (Schriftcode: Hant) 珍·奧斯婷 (Chinesische Namensform (Transkription um 1950)) (Schriftcode: Hant) 简·奧斯汀 (Chinesische Namensform (Transkription wikipedia, Zugriff: 18.05.2021)) (Schriftcode: Hans) 簡·奧斯汀 (Chinesische Namensform (Transkription wikipedia, Zugriff: 18.05.2021)) (Schriftcode: Hant)
Quelle	M B 1986 Kröner Weltlit. LCAuth Wikipedia (Stand: 09.03.2023): https://de.wikipedia.org/w/index.php?title=Jane_Austen&oldid=231601159
Zeit	Lebensdaten: 1775-1817
Land	Großbritannien (XA-GB)
Sprache(n)	Englisch (eng)
Geografischer Bezug	Geburtsort: Steventon (Hampshire) Sterbeort: Winchester Wirkungsort: Bath
Beruf(e)	Schriftstellerin Librettistin
Beziehungen zu Personen	Austen, Cassandra (Schwester) Austen, James (Bruder) Austen, Charles (Bruder) De Feuillede, Eliza (Cousine) Hawke, Cassandra (Cousine) Knight, Caroline Jane (Großnichte) Austen-Leigh, Lois (Groß-Großnichte)

Fonte: Biblioteca Nacional da Alemanha (2025).

¹⁵ Catálogo de autoridades da DNB:

<https://portal.dnb.de/opac/opacPresentation?cqlMode=true&reset=true&referrerPosition=0&referrerResultId=%22AUSTEN%2C%22+and+%22JANE%22%26any&query=idn%3D118505173>.

Em contato com a biblioteca, foi informado que o campo 375 é utilizado e não é obrigatório. O campo "*Geschlecht*" (Gênero) está preenchido com "*weiblich*" (feminino), evidenciando a prática da biblioteca em registrar essa informação.

Além disso, inclui nomes variantes, provenientes de diferentes sistemas de escrita e transliterações, como formas chinesas e russas, além de informações sobre a identidade da autora, como datas (1775-1817), local de nascimento e morte, língua associada (inglês) e profissão (escritora e libretista).

A padronização para inserir o gênero segue os códigos da ISO 5218¹⁶, utiliza-se apenas "m" (masculino) e "f" (feminino), já que não existe uma opção específica para gênero diverso, pois os códigos "0" e "9" representam "desconhecido" ou "não aplicável" e não são considerados adequados. No entanto, o uso desse campo ainda está em discussão, pois não há uma definição clara para "diverso", e há debates sobre registrá-lo apenas para pessoas falecidas. Ao mesmo tempo, pesquisadores que utilizam o arquivo de autoridade defendem a inclusão dessa informação no campo 375.

O catálogo da biblioteca adota técnicas da *Web Semântica*, permitindo a exportação dos registros em *RDF Turtle*. Essa abordagem insere os dados em um ambiente de *Linked Data*, favorecendo a interoperabilidade e conectividade entre diferentes bases. Essa integração demonstra a adaptação dos catálogos aos ambientes semânticos, possibilitando a representação e recuperação da informação mais contextualizada. Uma parte do mesmo registro em MARC 21 XML pode ser visualizado na Figura 12, evidenciando a estruturação dos dados conforme o padrão.

¹⁶ ISO 5218: https://pt.wikipedia.org/wiki/ISO/IEC_5218.

Figura 12 - Registro MARC 21 XML de autoridade da autora Jane Austen no catálogo da Biblioteca Nacional da Alemanha

```

<datafield tag="100" ind1="1" ind2=" ">
  <subfield code="a">Austen, Jane</subfield>
  <subfield code="d">1775-1817</subfield>
</datafield>
<datafield tag="375" ind1=" " ind2=" ">
  <subfield code="a">2</subfield>
  <subfield code="2">iso5218</subfield>
</datafield>
<datafield tag="377" ind1=" " ind2="7">
  <subfield code="a">eng</subfield>
  <subfield code="2">iso639-2b</subfield>
</datafield>
<datafield tag="400" ind1="0" ind2=" ">
  <subfield code="a">Aosiding</subfield>
  <subfield code="d">1775-1817</subfield>
  <subfield code="9">v:Chines.</subfield>
</datafield>
<datafield tag="400" ind1="1" ind2=" ">
  <subfield code="a">Aosiding, Jian</subfield>
  <subfield code="d">1775-1817</subfield>
</datafield>
<datafield tag="400" ind1="1" ind2=" ">
  <subfield code="a">Ao-ssu-ting, Chien</subfield>
  <subfield code="d">1775-1817</subfield>
</datafield>
<datafield tag="400" ind1="0" ind2=" ">

```

Fonte: Biblioteca Nacional da Alemanha (2025).

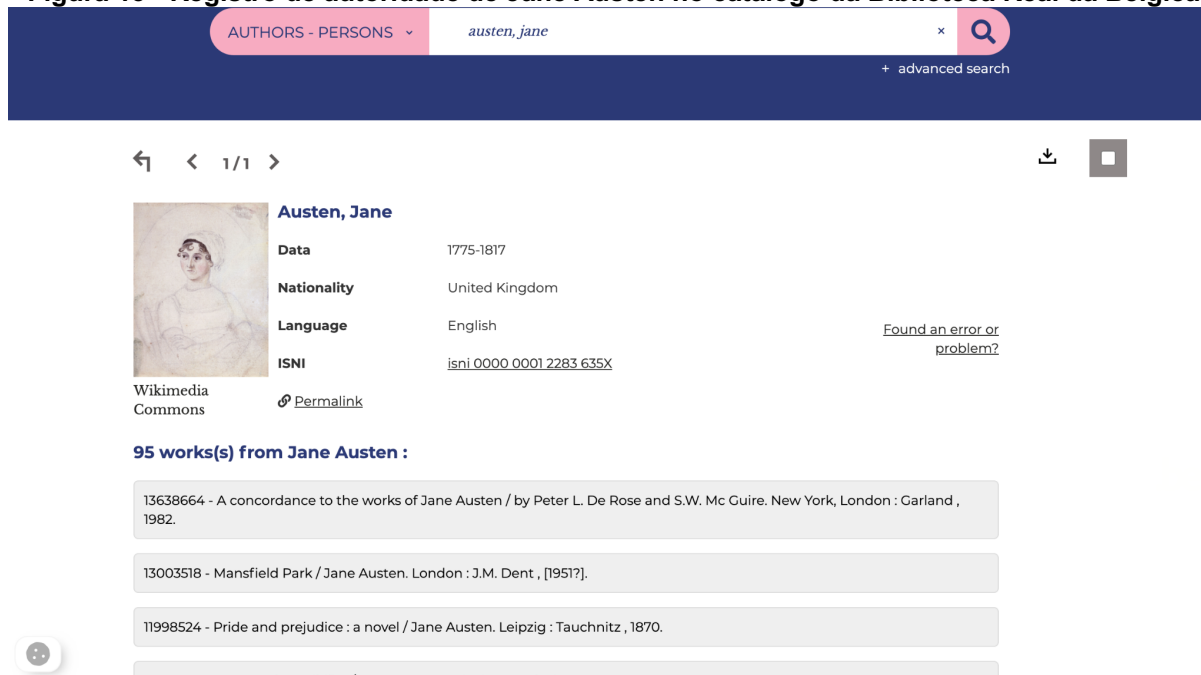
Como demonstrado na Figura 12, o campo 375 aparece codificado com o valor “2” no subcampo \$a, e vinculado ao vocabulário de origem no subcampo \$2 com a designação “iso5218”, indicando conformidade com a norma ISO 5218, na qual o código “2” corresponde ao gênero feminino.

4.2.6 Biblioteca Real da Bélgica

Na Figura 13, está o registro de autoridade de *Jane Austen* na Biblioteca Real da Bélgica (KBR), no catálogo de autoridade¹⁷.

¹⁷ Catalogo KBR:

Figura 13 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Real da Bélgica



AUTHORS - PERSONS

+ advanced search

← < 1/1 >

Austen, Jane

Data	1775-1817
Nationality	United Kingdom
Language	English
ISNI	isni 0000 0001 2283 635X

[Found an error or problem?](#)

Wikimedia Commons

95 works(s) from Jane Austen :

- 13638664 - A concordance to the works of Jane Austen / by Peter L. De Rose and S.W. Mc Guire. New York, London : Garland , 1982.
- 13003518 - Mansfield Park / Jane Austen. London : J.M. Dent , [1951?].
- 11998524 - Pride and prejudice : a novel / Jane Austen. Leipzig : Tauchnitz , 1870.
- 13003518 - Mansfield Park / Jane Austen. London : J.M. Dent , [1951?].

Fonte: Biblioteca Real da Bélgica (2025).

São exibidas informações como datas de vida (1775-1817), nacionalidade (Reino Unido), idioma (inglês) e ISNI (0000 0001 2283 635X). Além disso, há uma lista de obras associadas à autora. Existe a integração com ferramentas da *Web Semântica*, permitindo a conexão dos dados com a *Wikimedia Commons* e o ISNI, o que viabiliza um ambiente em *Linked Data*. Essa abordagem permite a interoperabilidade entre diferentes bases de dados e possibilita a recuperação e enriquecimento dos metadados de maneira estruturada.

No entanto, não foi possível visualizar o formato MARC 21 no catálogo. Em contato com a KBR, foi informado que o campo 375 (Gênero) é utilizado nos registros de autoridade e que o catálogo permite a aplicação de filtros por gênero a partir desse dado, conforme a Figura 14.

Figura 14 - Separação por gênero no catálogo da Biblioteca Real da Bélgica

The screenshot shows a search interface with a search bar containing 'austen, jane' and a filter dropdown set to 'AUTHORS - PERSONS'. The search results are displayed in two sections. The first section, 'Search results', shows 'My search: austen, jane' and 'Personne' as filters. It lists 'Results 1 - 1 / 1' and 'Sort: Relevance'. The result is 'Austen, Jane 1775-1817' with the authority type 'Personne physique'. The second section, 'Refine your search', lists filters for 'Nationality' (United Kingdom), 'Sexe' (Female), 'Language' (English), and 'Type of authority' (Personne, Personne physique).

Fonte: Biblioteca Real da Bélgica (2025).

O filtro é baseado no campo 375 e as diretrizes de catalogação para esse campo são básicas e seguem os padrões do RDA. Os códigos utilizados são: M (male - masculino), F (female - feminino), U (unknown - desconhecido) e X (not applicable - não aplicável). A integração dessa informação ao catálogo da KBR permite ampliar a capacidade de organização e recuperação da informação, promovendo a visibilidade de autoras e autores por meio de filtros acessíveis ao público. Trata-se de um exemplo de como a utilização estruturada do metadado de gênero pode ser funcionalmente explorada para fins de descrição, acesso e representação mais inclusiva nos sistemas bibliográficos.

4.2.7 Biblioteca Nacional da Espanha

Na figura 15 está a busca feita a partir de “Austen, Jane” na *Biblioteca Nacional de España* (BNE).

Figura 15 - Registro MARC 21 de autoridade da autora Jane Austen no catálogo da Biblioteca Nacional de Espanha

```

leader 01128cz za2200289n 4500
001 981060881094408606
005 20240801122906.0
008 901120nn aijnnaabn n aaa
016 ##$aBNE19900162256
024 7#$ahttp://viaf.org/viaf/102333412/ $2viaf
024 8#$ahttps://data.cervantesvirtual.com/person/471
024 7#$ahttps://isni.org/isni/000000012283635X $2isni
024 7#$ahttp://www.wikidata.org/entity/Q36322 $2wikidata
035 ##$a(SpMaBN)XX1124986
040 ##$aSpMaBN $bspa $cSpMaBN $serdc $fembne
046 ##$f1775 $g1817
100 1#$aAusten, Jane $d1775-1817
370 ##$aSteventon $vWikidata
372 ##$aFicción gótica $aNovela romántica $vWikidata
374 ##$aNovelistas $aEscritores $vWikidata
375 ##$aFemenino ←
377 ##$lInglés
670 ##$aSentido y sensibilidad, 2014 $bport. (Jane Austen)
670 ##$aOrgullo y prejuicio, 1986 $bport. (Jane Austen)
670 ##$aWW BL, 2-4-2014 $b(Austen, Jane; 1775-1817)
670 ##$aWW Wikipedia, 2-4-2014 $b(Jane Austen; n. 16-12-1775 en Steventon-m. 18-7-1817, Winchester; novelista británica contada entre los
935 ##$aXX1124986

```



Aviso legal

Fonte: Biblioteca Nacional de Espanha (2025).

O campo 375 foi utilizado para indicar o gênero da autora *Jane Austen*, com o valor registrado como “Femenino” (feminino). A presença desse campo evidencia que a instituição adota a prática de registrar o gênero em seus registros de autoridade, o que possibilita a identificação de autoras no catálogo com base em atributos descritivos mais específicos.

No campo 100 o nome autorizado da autora acompanhado de suas datas de nascimento e falecimento: *Austen, Jane, 1775-1817*. O campo 046 também especifica essas datas de forma codificada, e o campo 372 informa as áreas de atividade intelectual da autora, com destaque para “ficção gótica” e “romance”. O campo 374 descreve as funções profissionais desempenhadas, indicando “romancista” e “escritora”.

O campo 377 identifica o idioma vinculado à produção da autora como o inglês. Já o campo 670 reúne as fontes de informação utilizadas para a construção do registro, como a *Wikipedia* e a Biblioteca Nacional Britânica. Por fim, o campo 024 apresenta identificadores externos como VIAF, ISNI e *Wikidata*, caracterizando o uso de técnicas semânticas. A adoção desses identificadores fortalece a

¹⁸ Registro no catálogo da BNE:

https://catalogo.bne.es/discovery/authfulldisplay?docid=981060881094408606&context=L&vid=34BNE_INST:CATALOGO&lang=es&offset=10.

interoperabilidade do registro com outras bases de dados na *Web Semântica*, ampliando a visibilidade, a conectividade e reuso dos dados.

Na Figura 16, a BNE, possui uma iniciativa em “*Autores en la BNE*”¹⁹, onde categoriza autores e autoras separadamente, além de classificá-los por áreas como médicos, artistas, bibliotecários, músicos, jornalistas, escritores etc.



Fonte: Biblioteca Nacional de España (2025).

A BNE informou por e-mail que realizou uma análise retrospectiva dos registros de autoridade para a marcação massiva do campo 375 com os valores "Masculino" e "Femenino", e não foram considerados outros gêneros. Desde então, os catalogadores passaram a incluir essa informação regularmente nos novos registros. O tratamento foi automatizado e incluiu estratégias como a identificação de termos que podem indicar o gênero nos registros (como "*papa*", "*rey*", "*condesa*"), a comparação dos nomes próprios com listas do Instituto Nacional de Estatística da Espanha, a integração de dados da VIAF para autores estrangeiros e o cruzamento de informações com o *Wikidata*. Esse processo de integração com fontes externas se destaca com um processo de enriquecimento semântico bastante utilizado nas instituições LAM.

¹⁹ Autores en la BNE: <https://www.bne.es/es/autores/listado-autores>.

4.2.8 Biblioteca Nacional da República Checa

Na Figura 17, está o registro de autoridade da autora *Jane Austen* na Biblioteca Nacional da República Checa²⁰.

Figura 17 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional da República Checa

Choose format: [Standard](#) -- [MARC](#)

Record 1 out of 18

LDR	-----cz--a22-----n--4500
FMT	JA
001	jn19990000321
003	CZ-PrNK
005	20161213082817.0
008	000622 n aznnaabn----- a a-----
040	a ABA001 b cze d ABA001 d HBG001 e rda
046	f 1775 g 1817
1001	a Austen, Jane, d 1775-1817 7 n19990000321
370	c Velká Británie f Anglie
372	a literatura
374	a spisovatelky
375	a žena ←
377	a eng
4001	a Osten, Džejn, d 1775-1817 2 czerus 0 o
670	a NKC
670	a Genijeva, Je. Ju.: Džejn Osten
670	a Rozum a cit
670	a www(Wikipedie, otevřená encyklopedie), cit. 13. 12. 2016
670	a LC (Names), cit. 7. 12. 2016 b autoritní forma, data, biografické údaje
6780	a Anglická spisovatelka.
85642	u https://cs.wikipedia.org/wiki/Jane_Austenov%C3%A1 4 N
906	a im19990101
906	a no20030127 b svejja
906	a op20011009 b glob c Zdroj
906	a op20030127 b jazi c Viz - nově; Zdroj; Biografická poznámka - nově
906	a op20030901 b anmu c Viz - další výskyt
906	a no20161212 b xsuti c Zdroj - další výskyt; pole 370, 372, 374, 375, 377 - nově
906	a op20161213 b tone c Viz - zrušeno (přechýlený tvar), zdroj - další výskyt, URL
SYS	000053066

[https://aleph.nkp.cz/F/?func=direct&doc_number=000053066&local_base=AUT]

Fonte: Biblioteca Nacional da República Checa (2025).

Observa-se a presença do campo MARC 375, utilizado para a indicação do gênero, preenchido com o termo *žena* (mulher). Em 2015, o RDA Original foi implementado na biblioteca e de acordo com as diretrizes adotadas os registros de autoridade podem conter o campo 375, quando aplicável, sendo este um elemento recomendado, mas não obrigatório. Para preenchimento desse campo, o sistema utiliza duas opções controladas, *muž* (homem) ou *žena* (mulher).

O registro inclui outros campos do bloco 3XX, como o 372, que indica a área de atividade (*literatura*), e o 374, que informa a ocupação (escritora). Esses elementos,

²⁰ Biblioteca Nacional da República Checa:

https://aleph.nkp.cz/F/6TMRJH13R2GM111BFDEMMFFPN9EBRGVVI8IEFEUR985J9IKC6E-29148?func=full-set-set&set_number=262853&set_entry=000001&format=001.

somados ao campo de gênero, ampliam o contexto descritivo da entidade, possibilitando uma representação mais precisa.

Observa-se ainda a inclusão de um identificador externo no campo 856, com uma URI para o registro correspondente na *Wikidata*, o que contribui para a interoperabilidade dos dados com outros sistemas e fontes abertas. Ainda que essa URI não esteja diretamente vinculada ao campo de gênero, sua presença reforça o potencial de vinculação entre os registros locais e fontes externas multilíngues e estruturadas, como preconizado nas práticas de dados vinculados.

4.2.9 Biblioteca Nacional da Islândia

Na sequência, a Figura 18 apresenta o registro de autoridade da autora J. L. Butler no catálogo da Biblioteca Nacional da Islândia²¹.

Figura 18 - Registro de autoridade de J. L. Butler no catálogo da Biblioteca Nacional da Islândia

PERSONAL NAME
Butler, J. L.
Search for records with this term >

TOP
SEND TO
DETAILS
LINKS

Send to

EMAIL PRINT PERMALINK

Details

Heading	Butler, J. L.
Related terms	Perry, Tasmina. >
Source data found	Heimild: Mine, 2019
Source	981014194438006886
Associated Language	eng

Links

Display Source Record >

Fonte: Biblioteca Nacional da Islândia (2025).

²¹ Biblioteca Nacional da Islândia:

https://lbs.primo.exlibrisgroup.com/discovery/authfulldisplay?docid=981014194438006886&context=L&vid=354ILC_NETWORK:10000_UNION&lang=en&offset=0.

O nome autorizado é apresentado como “Butler, J. L.”, com referência explícita à autora *Tamsin Perry* como identidade associada, e idioma vinculado como “inglês” (eng).

Figura 19 - Registro MARC 21 de autoridade de J. L. Butler no catálogo da Biblioteca Nacional da Islândia

```

leader 00277nz a2200121n 4500
001 981014194438006886
005 20230116105859.0
008 190528|a|aczn|aab|||||||||||aaa||| c
100 1#$aButler, J. L.
375 ##$akona ←
377 ##$aeng
500 1#$aPerry, Tasmia
670 ##$aHeimild: Mine, 2019

```

Fonte: Biblioteca Nacional da Islândia (2025).

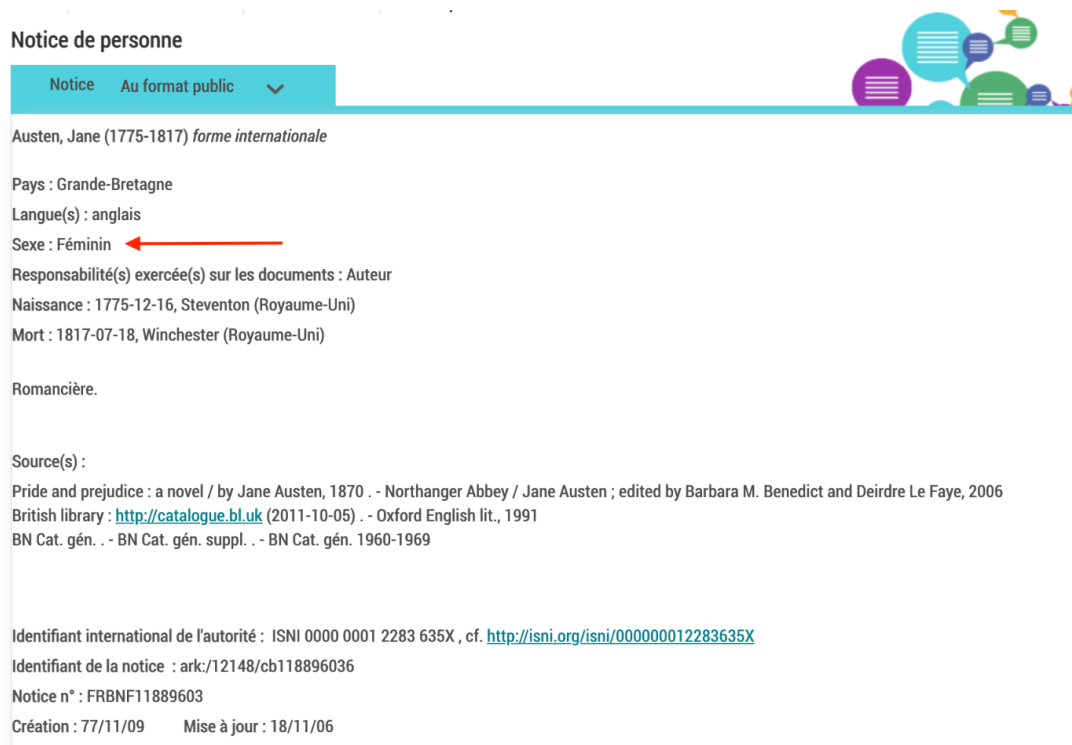
Na Figura 19, o registro de autoridade no formato MARC 21 J. L. Butler, no qual se faz o destaque para o uso do campo 375, que registra o gênero da entidade. O valor atribuído ao campo é *kona*, termo que significa "mulher" em islandês. Em contato com a biblioteca foi informado que seu uso não é obrigatório, e os valores utilizados são masculino, feminino e outros. Quando há dúvida quanto ao valor a ser atribuído, a recomendação é não incluir o campo no registro.

O registro inclui o campo 377, com os códigos de idioma *eng* (inglês). O campo 100 é a forma autorizada do nome que corresponde a um pseudônimo da autora Tasmia Perry, identificada no campo 500 como nome relacionado.

4.2.10 Biblioteca Nacional da França

Na Biblioteca Nacional da França (BnF), dois formatos são utilizados, na Figura 20 está o registro na interface pública da autora Jane Austen.

Figura 20 - Registro de autoridade de Jane Austen no catálogo da Biblioteca Nacional da França



Notice de personne

Notice Au format public

Austen, Jane (1775-1817) *forme internationale*

Pays : Grande-Bretagne

Langue(s) : anglais

Sexe : Féminin

Responsabilité(s) exercée(s) sur les documents : Auteur

Naissance : 1775-12-16, Steventon (Royaume-Uni)

Mort : 1817-07-18, Winchester (Royaume-Uni)

Romancière.

Source(s) :

Pride and prejudice : a novel / by Jane Austen, 1870 . - Northanger Abbey / Jane Austen ; edited by Barbara M. Benedict and Deirdre Le Faye, 2006

British library : <http://catalogue.bl.uk> (2011-10-05) . - Oxford English lit., 1991

BN Cat. gén. . - BN Cat. gén. suppl. . - BN Cat. gén. 1960-1969

Identifiant international de l'autorité : ISNI 0000 0001 2283 635X , cf. <http://isni.org/isni/000000012283635X>

Identifiant de la notice : ark:/12148/cb118896036

Notice n° : FRBNF11889603

Création : 77/11/09 Mise à jour : 18/11/06

Fonte: Biblioteca Nacional da França (2025).

Na interface pública da BnF, existe a informação sobre o gênero, “Sexe: Féminin”. Esse dado aparece entre os campos *Langue(s)* (idioma) e *Responsabilité(s) exercée(s)* (função exercida nos documentos), sendo claramente visível ao usuário. Além disso, o registro exhibe informações como país de origem *Grande-Bretagne* (Grã-Bretanha), idioma *anglais* (inglês), datas de nascimento e morte, ocupação (*Romancière*), fontes consultadas e identificadores internacionais como o ISNI. Na sequência, a Figura 21, está o registro no formato UNIMARC.

Figura 21 - Registro UNIMARC de autoridade de Jane Austen no catálogo da Biblioteca Nacional da França

Accueil > Recherche auteurs A-Z > Liste auteurs A-Z > Nom de personne

Notice de personne

Notice Au format Unimarc

```

000 cx a22 45
001 FRBNF11889603X
003 http://catalogue.bnf.fr/ark:/12148/cb118896036
005 20181106
010 .. $a 000000012283635X $2 VIAF $d 20130724
039 .. $o OPL $a 001573266 $z AMA
039 .. $o OPL $a 000419322 $z APP
039 .. $o OPP $a 14871547 $d 20070226
039 .. $o OPP $a 15447424 $d 20121008
100 .. $a 19771109afrey50 ba0
101 .. $a eng
102 .. $a GB
103 .. $a 17751216 18170718
105 .. $a a
106 .. $a 010
120 .. $a a
152 .. $c 2
200 | $7 ba0yba0y $8 fre $9 0 $a Austen $b Jane $f 1775-1817
300 | $a Romancière
301 .. $a Steventon (Royaume-Uni) $b Winchester (Royaume-Uni)
801 .. $a FR $b FR-751131015 $c 20181106
810 .. $a Pride and prejudice : a novel / by Jane Austen, 1870 $a Northanger Abbey / Jane Austen ; edited by Barbara M. Benedict and Deirdre Le Faye, 2006
810 .. $a British library - http://catalogue.bl.uk (2011-10-05) $a Oxford English lit., 1991
810 .. $a BN Cat. gén. $a BN Cat. gén. suppl. $a BN Cat. gén. 1960-1969
856 .. $u http://catalogue.bl.uk $e 2011-10-05

```

Fonte: Biblioteca Nacional da França (2025).

No formato UNIMARC, o gênero está no campo 120, subcampo \$a, com o valor “a”, que representa o feminino. O campo 100 identifica o número de controle do registro. O campo 103 apresenta a data de nascimento (1775). O campo 102 informa o país associado à entidade (GB, Reino Unido) e o campo 101 o idioma vinculado (a para inglês). O campo 200 apresenta a forma preferida, *Austen, Jane*. O campo 300 indica a ocupação da pessoa, com o termo *Romancière* (romancista), e o campo 801 informa a agência que criou o registro, confirmando que foi elaborado na França. O campo 856, por sua vez, oferece um *link* para uma fonte digital utilizada na catalogação, neste caso, o catálogo da *British Library*, reforçando a prática de utilizar fontes para a construção do registro. Já os campos 810 e 820 listam outras fontes consultadas, como obras publicadas e catálogos anteriores. Na figura 22, está o registro no formato INTERMARC na BnF.

Figura 22 - Registro INTERMARC de autoridade de Jane Austen no catálogo da Biblioteca Nacional da França

Accueil > Nom de personne

Notice de personne

Notice Au format InterMarc ▼

```

000 00778c1 ap22000272 45
001 FRBNF11889603X
008 771109181106gbengf 17751216 18170718 a 010
031 .. $a 000000012283635X $2 VIAF $d 20130724
045 .. $a a
100 .. $w .0..b..... $a Austen $m Jane
600 .. $a Romancière
603 .. $a Steventon (Royaume-Uni) $b Winchester (Royaume-Uni)
610 .. $a Pride and prejudice : a novel / by Jane Austen, 1870 $a Northanger Abbey / Jane Austen ; edited by Barbara M. Benedict and Deirdre Le Faye, 2006
610 .. $a British library $u http://catalogue.bl.uk $d 2011-10-05 $a Oxford English lit., 1991
610 .. $a BN Cat. gén. $a BN Cat. gén. suppl. $a BN Cat. gén. 1960-1969
917 .. $o OPL $a 001573266 $b AMA
917 .. $o OPL $a 000419322 $b APP
996 .. $o OPP $a 14871547 $d 20070226
996 .. $o OPP $a 15447424 $d 20121008

```

Fonte: Biblioteca Nacional da França (2025).

O registro da autora Jane Austen, conforme apresentado na Figura 25, apresenta o gênero registrado no campo 008, posição 17, com o valor “f” (feminino). Observa-se a presença do campo 031, que traz o vínculo ao VIAF, indicando a adoção de técnicas semânticas no catálogo. Essa vinculação permite a interoperabilidade com outras bases de dados e reforça o alinhamento do sistema a práticas de *Linked Data*. A inclusão de identificadores persistentes e a normalização de pontos de acesso mostram que, mesmo em um formato nacional como o InterMarc, a BnF aplica mecanismos que favorecem o enriquecimento semântico dos registros de autoridade, especialmente por meio do uso de URIs e interconexões com fontes externas reconhecidas internacionalmente.

4.2.11 Biblioteca Nacional da Eslovênia

O registro apresentado na Figura 23 foi extraído da documentação técnica da Biblioteca Nacional da Eslovênia e refere-se à autora Agatha Christie.

Figura 23 - Registro CONMARC/A de autoridade de Agatha Christie na documentação técnica da Biblioteca Nacional da Eslovênia

100	uu	ba ceng gba
101	uu	aeng
120	uu	aa ba
200	u1	aChristie, bAgatha, f1890-1976 <i>(The entity is female and the name is differentiated by the addition of dates of birth and death.)</i>

Fonte: Biblioteca Nacional da Eslovênia (2019).

O campo 120, utilizado no formato COMARC/A, é composto por dois subcampos: o subcampo \$a, que registra o gênero da entidade, e o subcampo \$b, que indica se o nome é diferenciado ou não. No exemplo analisado, o subcampo 120\$a assume o valor “a”, representando o gênero feminino, enquanto o subcampo 120\$b também recebe o valor “a”, indicando que o nome está diferenciado, ou seja, foi desambiguado com o acréscimo de datas de nascimento e morte no campo 200.

O campo 200 apresenta o ponto de acesso autorizado, com a forma: *Christie, Agatha, f1890–1976*. Esse exemplo ilustra como o campo 120 do COMARC/A permite a representação codificada e estruturada da identidade de gênero, bem como a distinção entre nomes homônimos, garantindo precisão e controle na construção dos registros de autoridade. A nota explicativa do exemplo reforça: *“The entity is female and the name is differentiated by the addition of dates of birth and death.”*, que em português seria “A entidade é do gênero feminino e o nome é diferenciado pela adição das datas de nascimento e falecimento.” Não foi possível verificar a existência de técnicas semânticas no catálogo.

4.2.12 Catálogo de autoridades de nomes e títulos da Catalunha

O Catálogo de autoridades CANTIC (Catàleg d’Autoritats de Noms i Títols de Catalunya) foi elaborado de forma cooperativa e é liderado pela Biblioteca da Catalunha. Na Figura 24, está o registro de autoridade²² de *Lola Anglada*, que é uma autora espanhola.

²²Registro de autoridade no CANTIC: <https://cantic.bnc.cat/detall/421957/26547>.

Figura 24 - Registro MARC 21 de autoridade de Lola Anglada no Catálogo de autoridades CANTIC

Etiqueta	1	2	Contingut
Capçalera			02537cz a2200469n 4500
001			981058507605406706
005			20241115002417.0
008			960307 azznnaabn a aaa
024	7		0000000071033230\$2isni\$1https://isni.org/isni/0000000071033230
024	7		Q459343\$2wikidata\$1http://www.wikidata.org/entity/Q459343
024	7		n80010244\$2lccn\$0http://id.loc.gov/authorities/names/n80010244
024	7		57522030\$2viaf\$1http://viaf.org/viaf/57522030
035			0777-17480\$9ES-BaCBU035
035			(ES-BaCBU).a10040663
040			ES-BaBC\$bcat\$erda\$cES-BaBC\$dES-BaBC\$dES-BaUB
046			\$f1892-10-28\$g1984-09-12\$2edtf
053		4	NC987.5.A64\$5ES-BaUPF
100	1		Anglada, Lola,\$d1892-1984
370			Barcelona (Catalunya)\$bTiana (Catalunya)\$cEspanya\$fCatalunya\$2cantic
372			Arts visuals\$2thub
374			Dibuixants\$aEscriptors\$2lemac
374			Il·lustradores\$aArtistes\$2thub
375			femení ←
400	1		Anglada, L.\$q(Lola),\$d1892-1984
400	1		Anglada i Sarriera, Lola,\$d1892-1984
400	1		\$wnna\$aAnglada-Sarriera, Lola,\$d1892-1984\$5ES-BaCBU

Fonte: Catálogo de autoridades CANTIC (2025).

O registro de autoridade apresentado, inclui o campo 375, que está preenchido com o valor "femení", indicando que a entidade descrita se identifica com o gênero feminino. Além disso, observa-se a presença dos campos 370, que informa o local de nascimento e a afiliação geográfica (Barcelona, Tiana, Espanha); 372, que registra a área de atividade da autora (Arts visuals); e 374, que identifica a ocupação (Il·lustradores). O campo 046 também está incluído, codificando as datas de nascimento e morte da entidade, e o campo 400 apresenta formas variantes do nome autorizado.

Com relação ao uso de técnicas semânticas, destaca-se a adoção do campo 024 com URI vinculada ao identificador ISNI, VIAF, *Wikidata* e a LC, evidenciando a integração com fontes externas. Essa vinculação semântica amplia a capacidade de conectividade com outras bases, promovendo o alinhamento com os princípios da *Web Semântica* e favorecendo a reutilização dos dados em contextos bibliográficos e não bibliográficos. Assim, o registro exemplifica uma abordagem estruturada e

semântica para representação de entidades, fortalecendo a qualidade e o alcance informacional dos dados de autoridade no CANTIC.

4.2.13 Biblioteca Nacional da Noruega

Por fim, na Figura 25 é apresentado o registro de autoridade de *Jane Austen* na Biblioteca Nacional da Noruega²³.

Figura 25 - Registro de autoridade na interface pública da autora Jane Austen no Biblioteca Nacional da Noruega

TIL TOPPEN

SEND TIL

DETALJER

ANDRE RESSURS...

Sitér

E-post

Skriv ut

Permalenke

Detaljer ^

Foretrukket navneform	Austen, Jane, 1775-1817
Type	Person
Navnevarianter	Osten, Džejn
Identifikatorer	VIAF: https://viaf.org/viaf/102333412 ISNI: https://isni.org/isni/000000012283635X
Lokal autoritets-ID	90068190
Kilde	BIBSYS autoritetsregister
Kvalitetsnivå	Kat3

Andre ressurser ^

MARC-visning (MARC21) >

Søk etter dokumenter som bruker denne autoriteten >

Felles autoritetsregister >

VIAF >

lln25 >

Fonte: Biblioteca Nacional da Noruega (2025).

Na Figura 25 o sistema exhibe informações básicas como a forma autorizada do nome (*Austen, Jane, 1775–1817*), a variante (*Osten, Džejn*), e os identificadores externos, incluindo links para os registros no VIAF e ISNI. Também são indicados o número de controle local (*90068190*), a fonte de origem do registro (BIBSYS), e o nível de qualidade catalográfica (*Kat3*). Embora o campo 375 não seja exibido diretamente na interface pública, ele está presente no registro na visualização MARC 21 que está na Figura 26.

²³ Catalogo de autoridade BIBSYS: https://bibsyst-almaprmo.hosted.exlibrisgroup.com/primo-explore/fulldisplay?docid=AUTREG90068190&context=L&vid=AUTREG&lang=no_NO&search_scope=default_scope&adaptor=Local%20Search%20Engine&tab=default_tab&query=any,contains,austen,%20jane&offset=0.

Figura 26 - Registro MARC 21 de autoridade de Jane Austen no catálogo da Biblioteca Nacional da Noruega

Marc-visning Sik

Austen, Jane (1775-1817)

Velg format: XML Linjeformat

Marc-post:

```
*ldr 99999nz a229999n 4500
*001 90068190
*003 NO-TrBIB
*005 20240415085508.0
*008 150506n| adz|naabn| |a|ana|
*0247# $ax90068190 $2NO-TrBIB
*0247# $ahttp://hdl.handle.net/11250/1305814 $2hdl
*0247# $ahttps://viaf.org/viaf/102333412 $2viaf
*0247# $ahttps://isni.org/isni/000000012283635X $2isni
*0247# $ahttps://id.bs.no/bibbi/72843 $2bibbi
*040# $aN0-TrBIB $bnob $cN0-TrBIB $fnoraf
*043# $agb
*1001# $aAusten, Jane $d1775-1817
*375# $af ←
*386# $aeng. $mNasjonalitet/regional gruppe $2bs-nasj
*4001# $aOsten, Dzejn
*901# $akat3
```

Last ned XML Last ned Linjeformat

Fonte: Biblioteca Nacional da Noruega (2025).

O campo 375 preenchido com o subcampo \$a contendo o valor “f”, indica que a entidade descrita se identifica com o gênero feminino. A descrição segue o formato MARC 21 e inclui o ponto de acesso autorizado no campo 100 (Austen, Jane \$d1775–1817), além do campo 386, que informa a nacionalidade da autora. O campo 400 traz uma variante nominal (Osten, Dzejn), evidenciando preocupação com a representação de diferentes formas de nome. Os campos 0247 incluem links para identificadores externos como VIAF e ISNI, o que demonstra a adoção de práticas semânticas e a vinculação a fontes de autoridade. Por fim, o campo 040 identifica a instituição catalogadora como NO-TrBIB, responsável pela criação e manutenção do registro. A estrutura revela o uso de metadados interoperáveis e alinhados com princípios da *Web Semântica*, promovendo a conectividade entre catálogos e a contextualização ampliada das entidades descritas.

A partir dos registros apresentados, no Quadro 10, são sintetizadas as informações relativas às bibliotecas analisadas, incluindo o formato de metadado, campo utilizado para registrar o metadado de gênero, os valores empregados e a presença de técnicas semânticas. Estas técnicas referem-se, sobretudo, à adoção de identificadores persistentes com URI e à incorporação de fontes externas vinculadas aos registros, como VIAF, ISNI e *Wikidata*. Ressalta-se que os registros apresentados nesta análise se referem exclusivamente a representação de entidades que se

identificam com o gênero feminino em registros de autoridade, em consonância com a linha de pesquisa da autora.

Quadro 11 - Bibliotecas que utilizam o metadado de gênero e técnicas semânticas

Biblioteca	Formato e campo utilizado	Valor	Técnicas Semânticas
Biblioteca Nacional Mariano Moreno	Campo 375 subcampo \$a do MARC 21	mujer	Campo 024 com VIAF e URI (via \$1).
Biblioteca Nacional da Colômbia	Campo 375 subcampo \$a do MARC 21	femenino	URI externa.
Biblioteca Nacional do México	Campo 375 subcampo \$a do MARC 21	femenino	Não identificado foi uso de técnicas semânticas.
Biblioteca Nacional do Peru	Campo 375 subcampo \$a do MARC 21	femenino	Integração com VIAF, ISNI e Wikidata.
Biblioteca Nacional da Alemanha	Campo 375 subcampo \$a do MARC 21	<i>weiblich</i>	RDF/XML; integração com LOD.
Biblioteca Real da Bélgica	Campo 375 subcampo \$a do MARC 21	f	Integração <i>Wikimedia Commons</i> e o ISNI.
Biblioteca Nacional de Espanha	Campo 375 subcampo \$a do MARC 21	femenino	Campo 024 com identificadores externos
Biblioteca Nacional da República Checa	Campo 375 subcampo \$a do MARC 21	<i>žena</i>	Campo 856 vinculado à Wikidata.
Biblioteca Nacional da Islândia	Campo 375 subcampo \$a do MARC 21	kona	Não foi identificado uso de técnicas semânticas.
Biblioteca Nacional da França	INTERMARC Campo 17 008 UNIMARC 120\$a	f a	Integração com VIAF.
Biblioteca Nacional da Eslovênia	Campo 120 subcampo \$a do CONMARC/A	a	Não foi identificado uso de técnicas semânticas.
Catálogo de autoridades CANTIC	Campo 375 subcampo \$a do MARC 21	femení	024 com URI vinculada ao identificador ISNI, VIAF, <i>Wikidata</i> e a LC.
Biblioteca Nacional da Noruega	Campo 375 subcampo \$a do MARC 21	f	0247 incluem links externos como VIAF e ISNI.

Fonte: elaborado pela autora (2025).

O Quadro 11 reúne de forma estruturada as informações extraídas da análise de registros de autoridade de bibliotecas nacionais que adotam o metadado de gênero. A primeira coluna, Biblioteca, identifica as instituições analisadas, destacando aquelas que disponibilizam catálogos com registros públicos e acessíveis para consulta. Todas as bibliotecas listadas incluíram, de alguma forma, o metadado de gênero em seus registros, permitindo a verificação direta ou por documentação técnica disponibilizada institucionalmente.

A segunda coluna, Formato e campo utilizado, apresenta os diferentes padrões de metadados empregados por cada biblioteca, com destaque para a predominância do formato MARC 21 e do campo 375, com uso do subcampo \$a. Esse campo foi criado para registrar o gênero da entidade nos registros de autoridade. Há, no entanto,

variações de implementação, a Biblioteca Nacional da França, por exemplo, utiliza o formato INTERMARC, com o valor de gênero codificado na posição 17 do campo 008, já a Biblioteca Nacional da Eslovênia utiliza o campo 120 do COMARC/A. Essas variações refletem a adoção de diferentes formatos de metadados.

A terceira coluna, Valores, identifica os termos ou códigos utilizados pelas instituições para indicar o gênero nos registros. Algumas bibliotecas, como a da Alemanha, Argentina, Colômbia, Espanha, México, Peru, República Checa, Islândia e o CANTIC adotam termos textuais completos para o gênero feminino, como *feminino*, *femení*, *weiblich*, *mujer*, *žena* e *kona*. Enquanto as bibliotecas da Bélgica, França, Eslovênia e Noruega seguem padrões codificados como F ou A.

Por fim, a quarta coluna, Enriquecimento semântico, apresenta se o registro apresenta evidências da adoção de técnicas semânticas. Foram considerados como indicadores de enriquecimento: o uso de identificadores persistentes com URI sendo parte dos princípios para a estruturação de dados em *Linked Data* segundo Berners-Lee (2006). A presença de links para fontes externas (como *Wikidata*, Biblioteca Nacional da França, Biblioteca Nacional da Espanha, entre outras) e o uso de campos específicos para vinculação a entidades externas, como o campo 024 ou 856. Bibliotecas como as da Argentina, Alemanha, França, Noruega, Colômbia, Espanha República Checa exemplificam abordagens que integram os registros a redes de dados interligados. Outras, como México, Islândia e Eslovênia, não evidenciam práticas de vinculação externa ou adoção de identificadores persistentes nos registros analisados.

Ainda que algumas bibliotecas façam uso parcial das boas práticas para a publicação de dados em *Linked Data*, o pleno aproveitamento das tecnologias semânticas ainda é uma lacuna. A ampliação do uso dessas técnicas pode fortalecer o enriquecimento semântico dos registros, potencializar a interoperabilidade, a precisão descritiva e a acessibilidade informacional, como destacado na Seção 3.

A partir dos resultados apresentados nos Quadros 10 e 11, as bibliotecas que utilizam o metadado de gênero o fazem predominantemente em sua forma binária, com os valores “masculino” e “feminino”. Como observa Possemato (2025), trata-se de uma discussão em andamento, inserida em um processo de transformação das práticas de catalogação, a incorporação de tecnologias da *Web*, pela modelagem de entidades e pela crescente demanda por descrições mais inclusivas.

A discussão sobre os metadados descritivos e a catalogação evidencia que essas práticas não se limitam a aspectos técnicos e operacionais. Em ambas as abordagens, as decisões institucionais influenciam diretamente a forma como a informação é representada e, conseqüentemente, como será acessada pelos usuários.

Reconhecendo essa responsabilidade, o *Cataloguing Code of Ethics* (2021) orienta que as decisões dos profissionais da área devem ter como base os princípios de justiça, precisão e equidade, reforçando o compromisso ético com a representação fiel dos recursos informacionais. Sardo (2019) destaca também que a ética na produção e manutenção de metadados é essencial para garantir um ambiente informacional justo, transparente e acessível.

O princípio da representação, conforme descrito pela IFLA (2016), orienta que as entidades sejam descritas conforme sua forma mais reconhecida ou autodesignada. Nos registros analisados, esse princípio é atendido. Porém, embora o campo 375 esteja presente, a maioria das bibliotecas utiliza termos binários (masculino/feminino) que podem não contemplar a diversidade de autodesignações de gênero. Esse dado confirma a observação de Possemato (2025), ao apontar que a catalogação ainda adota modelos normativos.

A normalização e a coerência, essenciais para a consistência e recuperação da informação, são evidenciadas na adoção do MARC 21 como padrão predominante e no uso do campo 375, especialmente no subcampo \$a, para registro do gênero. Em outros casos, como no uso do COMARC/A (Eslovênia) ou INTERMARC (França), há variações, mas que seguem a mesma lógica de organização normativa.

A interoperabilidade e a acessibilidade, princípios diretamente vinculados à *Web Semântica*, aparecem de forma mais evidente nos registros que apresentam enriquecimento semântico. Tais práticas são verificadas quando os registros contêm URIs persistentes, campos como 024 (identificadores externos) e 856 (ligações para fontes externas), e integração com bases como *Wikidata* e VIAF. Além disso, no MARC 21 há o campo subcampo \$2, que indica a fonte do vocabulário controlado, pode afetar a interoperabilidade e a precisão na interpretação dos valores, evidenciando a importância de sua utilização para garantir alinhamento semântico. Essas práticas são fundamentais para promover a reutilização e o compartilhamento em ambientes de *Linked Data*.

O controle de autoridade é um dos pilares da consistência nos catálogos. Ele assegura que nomes de pessoas, entidades coletivas, títulos uniformes e outros pontos de acesso sejam registrados de forma padronizada, permitindo a recuperação eficaz da informação. Essa prática é essencial quando se observa a presença de nomes homônimos ou variações linguísticas. Ao incorporar o metadado de gênero, as bibliotecas podem oferecer um ponto de diferenciação adicional. No entanto, os resultados apontam que a maioria das instituições ainda adota valores binários, o que limita a representatividade das identidades e reduz o potencial do controle de autoridade como mecanismo de visibilidade e inclusão.

Olson (2002) reforça que os catálogos não são neutros, ao selecionar termos, estruturar assuntos e definir pontos de acesso, os catalogadores possuem o controle sobre o que vai ser recuperado.

Barbosa (1978) complementa que o catálogo é um instrumento de comunicação, e sua eficácia está diretamente relacionada ao grau de normalização e consistência na descrição dos dados. Nesse sentido, os catálogos devem garantir que os usuários consigam localizar, identificar, selecionar e acessar os recursos de forma eficiente, como define a IFLA (2016).

Dessa forma, ao relacionar os princípios de catalogação, os fundamentos da Família FR, as práticas de controle de autoridade e o uso de metadados descritivos com os resultados obtidos na análise empírica, é possível observar avanços pontuais, mas também limitações significativas na forma como o gênero é representado nos registros de autoridade. A adoção predominante de modelos binários e o uso ainda incipiente de recursos semânticos indicam que o campo continua em transformação. A heterogeneidade observada entre os catálogos reflete tanto diferenças de ações institucionais quanto as tecnologias de cada biblioteca no que tange à adoção da *Web Semântica* e ao enriquecimento semântico dos registros.

Frente a esse cenário, reafirma-se a importância de iniciativas que busquem aprimorar a representação da identidade de gênero nos registros de autoridade, considerando a diversidade de identidades e a necessidade de garantir interoperabilidade e acessibilidade nos sistemas de informação. Considerando esses aspectos, a próxima seção evidencia um conjunto de boas práticas para o aprimoramento da representação da identidade de gênero em registros de autoridade. Essa proposição é fundamentada nas práticas observadas nas bibliotecas analisadas, nas instruções do RDA, nos Princípios Internacionais de Catalogação (IFLA, 2016) e

nas possibilidades oferecidas pelas tecnologias da *Web Semântica*. Mais do que sugerir diretrizes normativas, busca-se evidenciar caminhos possíveis e sustentáveis para qualificar os registros de autoridade a partir de uma perspectiva técnica, ética e informacional.

5. BOAS PRÁTICAS A DETERMINAÇÃO DO METADADO DE GÊNERO

A partir dos resultados da análise dos registros de autoridade apresentados na Seção 5 e da articulação com os fundamentos teóricos, esta seção evidencia boas práticas voltadas à melhoria da representação da identidade de gênero nos registros de autoridade. A proposta não busca estabelecer um padrão normativo, mas reunir recomendações baseadas em princípios éticos, técnicos e semânticos que fortaleçam a qualidade, a inclusão e a interoperabilidade dos dados.

Para a determinação do metadado de gênero é preciso revisitar os *Princípios Internacionais de Catalogação* (IFLA, 2016), que visam promover uma abordagem consistente e centrada na conveniência do usuário. Relembrando que os princípios são, interesse do usuário, uso comum, representação; precisão, suficiência e necessidade; significantes, economia, coerência e a normalização, integração, interoperabilidade, abertura, acessibilidade e racionalidade (IFLA, 2016). Além disso, fornecem uma base metodológica sólida para pensar a catalogação não apenas como atividade técnica, mas como prática informacional com implicações sociais, culturais e éticas.

Do ponto de vista técnico, a recomendação amplamente adotada é pelo uso do campo 375 do formato MARC 21, particularmente o subcampo \$a, destinado ao registro do termo relativo ao gênero em registros de autoridade de nome pessoal. Este campo integra o bloco de campos 3xx de atributos da entidade, sendo projetado para representar características como gênero, grupo étnico, nacionalidade, entre outros. A inclusão estruturada dessa informação facilita a recuperação, desambiguação e organização dos dados em catálogos e bases cooperativas.

Além do subcampo \$a, destaca-se a relevância do subcampo \$2 no MARC 21, que permite indicar a fonte do vocabulário controlado utilizado, como o *Library of Congress Demographic Group Terms* (LCDGT), entre outros. Essa informação semântica complementa o conteúdo registrado, promovendo uma interpretação precisa dos termos e favorecendo a interoperabilidade entre sistemas que operam sob diferentes padrões ou contextos linguísticos. O uso do subcampo \$2 permite, ainda, a adesão às boas práticas de catalogação orientadas pela *Web Semântica*, ao permitir a vinculação com fontes externas padronizadas e com identificadores persistentes.

A integração dos registros de autoridade a redes de dados interligados é outro eixo central dessas boas práticas. Isso pode ser feito por meio do uso dos campos 024 (Identificador Padrão) e 856 (Localização e Acesso Eletrônico), que permitem associar o registro a identificadores persistentes, como *Wikidata*, VIAF, ISNI, entre outros. A presença de URIs nos registros é considerada uma evidência de enriquecimento semântico, conforme os princípios do *Linked Data* propostos por Berners-Lee (2006), permitindo que os registros se tornem parte de uma rede de dados mais ampla, conectada e reutilizável por diferentes sistemas informacionais.

Em casos em que o gênero não é informado, alterado ao longo do tempo ou suscetível de interpretação, recomenda-se registrar notas explicativas nos campos 670 (Fonte Consultada) ou 678 (Nota Biográfica). Estas notas permitem fornecer contexto, evitando suposições indevidas, além de promover transparência na origem da informação registrada. Sendo coerente com os princípios de precisão, significação e necessidade, que orientam que apenas os elementos relevantes e devidamente verificados devem compor os registros de autoridade.

Contudo, nem todas as instituições utilizam exclusivamente o formato MARC 21. No caso da Biblioteca Nacional da França, que adota o formato INTERMARC, a informação de gênero é registrada no campo 120. Este campo abriga atributos pessoais como sexo e data de nascimento, sendo o gênero representado por códigos alfanuméricos definidos pela instituição. Embora o INTERMARC não possua a mesma granularidade de subcampos do MARC 21, ele cumpre função equivalente ao oferecer uma estrutura compatível para o registro desses dados.

Além das práticas técnicas, é fundamental que as instituições documentem e publiquem suas ações e políticas de catalogação, e explicitem em quais esquemas de descrição se basearam, como por exemplo, o RDA Original e Oficial, nos princípios da IFLA e nas recomendações internacionais sobre inclusão e ética na descrição de dados. Essas políticas devem orientar o uso responsável do campo de gênero e estabelecer parâmetros para a decisão sobre quando e como registrar tal informação. Também, se possível, caso optem por utilizar o metadado de gênero, devem abordar aspectos como a revisão de registros anteriores, a exclusão de informações imprecisas e o uso de linguagem neutra, quando apropriado.

A adoção de práticas éticas é reforçada pelo *Cataloguing Code of Ethics* (2021), que afirma o compromisso dos profissionais da informação com a justiça descritiva, a não discriminação e a representatividade. A determinação do gênero nos registros de

autoridade, portanto, não pode ser tratada apenas como uma tarefa técnica, mas como uma ação com implicações diretas sobre a visibilidade, o respeito e a legitimidade das identidades representadas nos sistemas de informação.

É importante destacar que os metadados descritivos, ao lado do controle de autoridade, exercem papel crucial na estruturação da memória institucional e cultural. Ao garantir a padronização e a consistência dos pontos de acesso, o controle de autoridade, conforme definido nos modelos da família FR (FRBR, FRAD, FRSAD) e consolidado no IFLA LRM, promove a unicidade e a integridade dos registros bibliográficos. No entanto, essa estrutura só será efetiva se for capaz de representar, de forma justa e adequada, a multiplicidade das identidades presentes nos acervos.

Portanto, as boas práticas aqui reunidas têm como objetivo orientar a atuação de profissionais e instituições na construção de registros mais inclusivos, éticos, conectados e interoperáveis, reconhecendo que a representação da identidade de gênero em registros de autoridade é, ao mesmo tempo, um desafio técnico e uma escolha política e ética.

A partir da análise dos registros de autoridade de bibliotecas nacionais e na revisão teórico-técnica desenvolvida ao longo desta pesquisa, identificaram-se fragilidades e lacunas na representação do metadado de gênero, bem como práticas pontuais que podem ser consideradas exemplares. A partir dessas evidências, foi possível elaborar um conjunto de diretrizes que buscam orientar a inclusão qualificada do metadado de gênero em registros de autoridade, respeitando princípios éticos, normas técnicas e exigências de interoperabilidade em ambientes digitais e semânticos.

As diretrizes a seguir foram organizadas em três dimensões complementares: técnica, ética e de interoperabilidade, conforme sintetizado no Quadro 12. Essa abordagem busca integrar os fundamentos normativos da catalogação com os compromissos sociais e tecnológicos necessários à construção de sistemas informacionais mais inclusivos, transparentes e conectados.

Quadro 12 – Diretrizes para a inclusão do metadado de gênero em registros de autoridade

Diretriz proposta	Fundamento técnico	Fundamento ético	Fundamento de interoperabilidade
1. Utilizar o campo 375 (MARC 21) ou 120 (UNIMARC) para representação estruturada do gênero	Utilização de campos padronizados e específicos conforme documentação	Reconhecimento institucional da identidade de gênero como atributo relevante para	Facilita a integração entre registros de diferentes instituições que adotam MARC 21 ou UNIMARC.

nos registros de autoridade.	técnica dos formatos de metadados.	representação informacional.	
2. Adotar vocabulários controlados sensíveis à diversidade de gênero , como GSSO ou termos definidos por fontes confiáveis.	Uso de vocabulários consistentes e codificados no subcampo \$2 do campo 375.	Evita abordagens binárias e respeita a autodeclaração.	Permite mapeamento entre esquemas e vocabulários em contextos multilinguísticos.
3. Priorizar a autodeclaração da identidade de gênero nas decisões catalográficas.	Fonte da informação documentada no subcampo \$v ou \$7, com grau de confiabilidade.	Respeito à identidade informacional da pessoa representada.	Garante consistência na atualização dos registros em redes cooperativas.
4. Vincular o campo 375 a URIs confiáveis , como Wikidata, VIAF ou bases de ontologias especializadas.	Inclusão do subcampo \$1 (URI de objeto do mundo real) e \$0 (número de controle).	Valoriza a transparência e a rastreabilidade da informação.	Fortalece a publicação dos registros como Linked Data e o uso em ambientes semânticos.
5. Evitar o uso de valores genéricos ou omissos , como “desconhecido” ou “não informado”, sempre que possível.	Evita ambiguidade nos processos de representação e recuperação.	Impede a invisibilização de identidades e promove justiça informacional.	Aumenta a granularidade e a expressividade dos registros em sistemas conectados.
6. Registrar o período de validade da informação de gênero , quando aplicável.	Uso dos subcampos \$s (data inicial) e \$t (data final).	Reconhecimento da fluidez da identidade de gênero ao longo do tempo.	Apoia versionamento e atualização dos dados em repositórios interoperáveis.
7. Adotar diretrizes institucionais claras e públicas para o uso do metadado de gênero.	Define critérios objetivos para o preenchimento do campo, minimizando inconsistências.	Garante transparência e equidade nas políticas descritivas.	Permite a harmonização entre registros cooperativos e redes de autoridades.

Fonte: elaborado pela autora (2025).

Essas diretrizes não pretendem estabelecer um modelo único ou prescritivo, mas sim indicar caminhos possíveis para a qualificação dos registros de autoridade no que se refere à identidade de gênero. Sua adoção requer o envolvimento das instituições responsáveis pelos catálogos, o desenvolvimento de políticas internas sensíveis à diversidade e o alinhamento com padrões internacionais de metadados e tecnologias de dados conectados. Ao integrar aspectos descritivos, éticos e técnicos, estas recomendações contribuem para ampliar a representatividade e a confiabilidade dos registros, promovendo uma catalogação mais justa, precisa e socialmente responsável.

6 CONSIDERAÇÕES FINAIS

A estruturação adequada dos metadados é decisiva para garantir a precisão, coerência e interoperabilidade dos registros em sistemas bibliográficos. Metadados são elementos fundamentais para a organização, descrição e recuperação da informação, e, no caso da representação da identidade de gênero que se apresenta como ponto focal desta pesquisa. Além disso, metadados bem estruturados permitem que as instituições avancem na direção de sistemas mais conectados e responsivos às demandas contemporâneas. O objetivo geral desta pesquisa consistiu em analisar a presença do metadado de gênero nos registros de autoridade dos catálogos das bibliotecas nacionais, identificando as decisões adotadas por essas instituições.

Desse modo, o primeiro objetivo específico, **analisar os fundamentos dos metadados e os padrões de metadados aplicados aos registros de autoridade, com ênfase nos modelos descritivos e estruturais utilizados**, foi atendido pela Seção 2, essa seção, foi discutida as classificações e tipologias dos metadados e os padrões de metadados, em especial o MARC 21 que possui o campo 375 para o registro do gênero.

Foram apresentados os modelos conceituais da família FR (FRBR, FRAD e FRSAD), fundamentais para a construção lógica das entidades, atributos e relacionamentos no contexto da catalogação, assim como a evolução para o IFLA LRM, que unifica esses modelos e fornece uma estrutura orientada à *Web Semântica*. O entendimento desses modelos é essencial para compreender como os registros de autoridade são organizados e como a representação da identidade de gênero, desafiam e expandem essas estruturas tradicionais. Também foi abordado como o RDA Original, enquanto esquema de descrição, operacionaliza esses modelos por meio da definição de atributos como “gênero” e da proposta de registros mais granulares, interoperáveis e preparados para ambientes em *Linked Data*.

E por fim, os fundamentos da catalogação e as abordagens relacionadas à inclusão do atributo gênero, evidenciando os debates em curso sobre a pertinência, os limites e os impactos dessa prática nos sistemas bibliográficos. Assim, essa fundamentação teórica sustentou as análises desenvolvidas nas seções seguintes do trabalho.

Na Seção 3, foi apresentado como o uso de tecnologias como RDF, SPARQL, URIs e ontologias podem ampliar a capacidade dos sistemas bibliográficos de

representar dados de forma mais contextualizada, conectada e interoperável. O estudo de casos institucionais evidenciou que o enriquecimento semântico permite integrar diferentes fontes de informação, melhorar a recuperação por meio de vínculos entre entidades e atributos. Essa abordagem se mostra estratégica para que bibliotecas, arquivos e museus possam se inserir nos ambientes de *Linked Data* e expandir o alcance e a reutilização de seus dados.

Os objetivos **mapear a presença do metadado de gênero nos registros de autoridade dos catálogos de bibliotecas nacionais e identificar diretrizes e políticas adotadas para a inclusão e padronização para a representação da identidade de gênero nos registros de autoridade** são complementares, sendo assim foram alcançados na Seção 4. Foram analisados os registros de autoridade de autoras de diversas bibliotecas nacionais e as diretrizes institucionais quanto à utilização do campo de gênero. A investigação mostrou diferentes graus de adesão ao uso do campo 375 do formato MARC 21 ou do campo 120 do UNIMARC, bem como distintas justificativas para sua utilização ou não, incluindo ações baseadas no RDA, preocupações com privacidade de dados pessoais e decisões internas. A análise evidenciou, ainda, que algumas instituições implementaram práticas automatizadas de marcação retroativa de gênero, enquanto outras seguem orientações do PCC de não utilizar e retirar a informação sobre o gênero dos registros. Dessa forma, o levantamento permitiu identificar tanto os procedimentos técnicos adotados quanto os fundamentos normativos e éticos que orientam a representação da identidade de gênero nos registros de autoridade em nível internacional.

Por fim, o último objetivo foi **propor a utilização de ferramentas semânticas para o aprimoramento da representação da identidade de gênero em registros de autoridade**, e foi alcançado na seção 5. Essa proposição foi construída com base nas potencialidades do uso de tecnologias da *Web Semântica*, nas práticas adotadas por bibliotecas nacionais e nas recomendações encontradas na literatura especializada.

Ressalta-se que essa ainda é uma discussão em andamento no campo da catalogação. Diversas instituições e redes cooperativas têm revisado suas ações, demandas sociais e reflexões sobre a função descritiva dos registros de autoridade. Por um lado, reconhece-se a importância da inclusão do gênero como elemento descritivo que pode contribuir para a visibilidade, recuperação e representatividade, especialmente de autoras e grupos historicamente invisibilizados. Por outro,

identificam-se preocupações com a privacidade das entidades descritas, sobretudo no caso de indivíduos vivos, bem como críticas à simplificação ou categorização inadequada de identidades diversas.

A representação da identidade de gênero em registros de autoridade não se limita a uma questão técnica ou descritiva, mas envolve decisões profundamente condicionadas por fatores culturais, históricos e institucionais. A presença ou a omissão do campo 375, destinado ao registro do gênero no formato MARC 21, reflete escolhas catalográficas que são, ao mesmo tempo, informacionais e políticas. Tais decisões estão imbricadas em visões de mundo que influenciam a forma como instituições documentais constroem e disseminam conhecimento. Em ambientes marcados por valores normativos tradicionais ou por concepções binárias de gênero, a resistência à inclusão desse campo pode estar associada à manutenção de modelos hegemônicos de representação, nos quais atributos identitários não conformes são invisibilizados, neutralizados ou sistematicamente ignorados.

Nesse contexto, observa-se que, mesmo em instituições que adotam padrões internacionalmente consolidados, como o MARC 21 ou o UNIMARC, a decisão de não utilizar o campo destinado ao gênero decorre, em grande medida, da ausência de políticas institucionais claras e da falta de consensos sobre os limites éticos e descritivos da inclusão desse dado. Argumentos como a preservação da privacidade, a pretensa neutralidade da catalogação ou a insegurança quanto aos termos a serem utilizados são frequentemente mobilizados para justificar tal omissão. No entanto, essas justificativas desconsideram que a não inclusão de atributos identitários nos registros bibliográficos reforça desigualdades estruturais, ao comprometer a visibilidade e a recuperação de obras produzidas por mulheres, pessoas trans e identidades não binárias. A escolha de não representar é, nesse caso, uma forma de representação silenciosa, que atua na reprodução de silenciamentos e na exclusão simbólica de determinados grupos nos sistemas de informação.

Adicionalmente, a adoção de atributos não binários ou estratégias baseadas na autodeclaração encontra resistências técnicas, normativas e culturais que revelam uma tensão entre a busca por padronização e a necessidade de garantir representações mais inclusivas. A ausência de vocabulários controlados atualizados, a rigidez de alguns esquemas descritivos e a dificuldade em integrar URIs provenientes de fontes abertas como Wikidata ou ontologias especializadas limitam a expressividade dos registros e restringem sua capacidade de refletir a pluralidade

identitária. Soma-se a isso o fato de que muitos sistemas legados de gestão bibliográfica não foram projetados para lidar com atributos flexíveis, contextuais ou autorreferenciados. Nesse cenário, torna-se necessário repensar as bases epistêmicas e éticas da catalogação, reconhecendo que a inclusão qualificada do campo 375 pode contribuir não apenas para a melhoria técnica dos registros, mas para a construção de catálogos mais justos, responsivos e alinhados aos princípios de equidade informacional e diversidade cognitiva.

Mas, ressalta-se que a inclusão qualificada do metadado de gênero em registros de autoridade oferece ganhos expressivos no que tange à recuperação da informação, à visibilidade de autorias e à diversidade representacional nos catálogos bibliográficos. Ao permitir a indexação e a filtragem por atributos identitários, o campo 375 amplia as possibilidades de análise e descoberta de obras, especialmente no que se refere à produção intelectual de mulheres, pessoas trans e outras identidades não hegemônicas. Essa ampliação tem impacto direto sobre estudos de gênero, análises bibliométricas, políticas de acervo e estratégias curatoriais voltadas à equidade. Em vez de operar apenas como um dado acessório, o metadado de gênero passa a integrar o conjunto de elementos descritivos que qualificam a representação da entidade e enriquecem semanticamente os sistemas de informação.

Sob a perspectiva da interoperabilidade, o uso do campo 375 com vocabulários controlados e URIs externas permite a vinculação dos registros de autoridade a redes abertas de dados, promovendo a integração entre diferentes instituições, domínios e plataformas. Isso favorece não apenas a consistência dos dados, mas também a reutilização e o enriquecimento automático de registros em contextos cooperativos e transnacionais. Ao incorporar elementos semânticos interoperáveis, como o subcampo \$1 (URI de objeto do mundo real), os registros se tornam mais expressivos e conectados, o que fortalece sua presença em ambientes digitais avançados, como bibliografias interligadas, catálogos cooperativos, motores de busca semânticos e sistemas de ciência aberta.

Além dos aspectos técnicos e informacionais, a presença do metadado de gênero nos registros de autoridade contribui para consolidar uma prática catalográfica socialmente responsável, alinhada aos princípios de justiça cognitiva e equidade informacional. Ao reconhecer e representar identidades diversas nos sistemas de memória, a catalogação passa a desempenhar um papel ativo na construção de ambientes mais inclusivos, sensíveis às transformações sociais e comprometidos com

a democratização do conhecimento. Trata-se, portanto, de um movimento que ultrapassa os limites da descrição bibliográfica tradicional, reposicionando os sistemas de informação como espaços de mediação cultural e representação simbólica, em que cada escolha descritiva carrega implicações éticas, políticas e epistêmicas.

Como desdobramento da presente pesquisa, propõe-se, em investigações futuras, a aplicação empírica da proposta de enriquecimento semântico em um catálogo de autoridade real, visando à validação prática dos princípios, diretrizes e técnicas discutidos ao longo deste estudo. Essa implementação permitiria avaliar a viabilidade técnica da incorporação do campo 375 com vocabulários controlados e URIs externas, bem como os impactos na recuperação da informação, na interoperabilidade e na visibilidade de identidades diversas. A experimentação em ambiente real configuraria um avanço significativo na consolidação de práticas mais inclusivas e conectadas à lógica dos dados abertos e interligados.

Paralelamente, identifica-se o desenvolvimento de um modelo de dados específico para a representação da identidade de gênero em registros de autoridade, alinhado aos princípios da *Web Semântica*, às ontologias de domínio e às normas internacionais de catalogação. Tal modelo poderia articular atributos descritivos, fontes de autodeclaração, vocabulários interoperáveis e mecanismos de atualização temporal, oferecendo uma estrutura flexível e eticamente fundamentada para o tratamento desse metadado. A proposta incluiria também a definição de classes, propriedades e relacionamentos capazes de expressar a complexidade e a fluidez das identidades de gênero nos sistemas informacionais.

Espera-se, com esta pesquisa, ter contribuído para a área da Ciência da Informação e da Biblioteconomia, ao analisar sobre como o metadado de gênero está sendo utilizado e se a informação sobre gênero é registrada em registros de autoridade.

REFERÊNCIAS

5 STARS OPEN DATA. **5 Stars Open Data**. 2015. Disponível em: <https://5stardata.info/en/>. Acesso em: 26 mar. 2025

ADOLPHO, K. **Who asked you? Consent, self-determination, and the report of the PCC Ad Hoc Task Group on Gender in Name Authority Records**. In: SANDBERG, Jane (ed.). *Ethical questions in name authority control*. Sacramento, CA: Library Juice Press, 2019. p. 111–131.

ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 132 f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2010. Disponível em: <https://repositorio.unesp.br/handle/11449/103361>. Acesso em: 26 mar. 2025.

ALVES, R. C. V. Metadados editoriais e livreiros: algumas considerações e relações com os padrões de metadados do domínio bibliográfico. **Encontro Nacional de Pesquisa em Ciência da Informação**, v. 19, p. 5205-5219, 2018.

ARAKAKI, A. C. S. O Modelo IFLA Library Reference Model e o Linked Data. **Informação & Informação**, Londrina, v. 25, n. 3, p. 163–186, 2020. Disponível em: <http://dx.doi.org/10.5433/1981-8920.2020v25n3p163>. Acesso em: 26 mar. 2025.

ARAKAKI, A. C. S.; ARAKAKI, F. A. Dados e metadados: conceitos e relações. **Ciência da Informação**, Brasília, DF, v. 49, n. 3, p. 34–45, set./dez. 2020.

ARAKAKI, F. A. **Linked data**: ligação de dados bibliográficos. 2016. 144 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, 2016. Disponível em: <https://repositorio.unesp.br/handle/11449/147979>. Acesso em: 26 mar. 2025.

ARAKAKI, F. A. **Metadados administrativos e a proveniência dos dados**: modelo baseado na família PROV. 2019. 139p. Tese (doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio Mesquita Filho”, Marília, 2019. Disponível em: <https://repositorio.unesp.br/handle/11449/180490>. Acesso em: 26 mar. 2025.

ARAKAKI, A. C. S.; ISOTANI, S. Metadados para a tomada de decisão baseada em evidências: proposta para a plataforma mec red. **Ciência da Informação**, v. 53, n. esp., 2024.

ARAKAKI, F. A.; SIMIONATO, A. C.; SANTOS, P. L. V. A. C. Catalogação e tecnologia: interseções com a web semântica. **Informação@Profissões**, v. 6, n. 2, p. 3-19, 2017. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/infoprof/article/view/32003>. Acesso em: 26 mar. 2025.

ARAKAKI, A. C. S.; ARAKAKI, F. A. **Resource Description and Access (RDA) oficial: aspectos e perspectivas para a catalogação**. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 2023. **Anais** [...] XXIII Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação, 2023.

ARAKAKI, F. A. *et al.* Disponibilização de metadados em linked data para repositórios digitais. In: **Anales de Documentación**. Facultad de Comunicación y Documentación y Servicio de Publicaciones de la Universidad de Murcia, 2019.

ASSUMPÇÃO, F. S. **Controle de autoridade**: definições, processo e componentes. Londrina: ABECIN, 2012. 122 p.

ASSUMPÇÃO, F. S. **Modelo para a publicação de dados de autoridade como Linked Data**. 208 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2018. Disponível em: <http://hdl.handle.net/11449/152759>. Acesso em: 26 mar. 2025.

ASSUMPÇÃO, F. S. **AACR2, MARC 21 e controle de autoridade**: um guia de estudo. 2020. p. 121-149. Disponível em: <https://fabricioassumpcao.com/blog/wp-content/uploads/2020/04/3-3-marc-21-de-autoridade.pdf>. Acesso em: 25 mar. 2025.

ASSUMPÇÃO, F. S.; PEREIRA, A. M.; TEIXEIRA, M. V. (org.). **RDA**: perspectivas teóricas e práticas no Brasil. Florianópolis: UDESC, 2020. ISBN: 978-65-88565-15-5. Disponível em: <https://rdanobrasil.org/wp-content/uploads/2021/05/rda-perspectivas-teoricas-e-praticas-no-brasil.pdf>. Acesso em: 25 mar. 2025.

BAKER, T. *et al.* **Library Linked Data Incubator Group Final Report**. W3C Incubator Group Report, 2011. Disponível em: <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>. Acesso em: 25 mar. 2025.

BARBOSA, A. P. **Novos rumos da catalogação**. Rio de Janeiro: BNG / Brasilart, 1978.

BERNERS-LEE, T. **Linked data**, 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 26 mar. 2025.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American**, may. 2001. Disponível em: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf. Acesso em: 26 mar. 2025.

BERTIN, M.; ATANASSOVA, I. Semantic enrichment of scientific publications and metadata. **D-lib Magazine**, v. 18, n. 7/8, 2012.

BILLEY, A.; DRABINSKI, E.; ROBERTO, K. R. What's gender got to do with it? A critique of RDA 9.7. **Cataloging & Classification Quarterly**, v. 52, n. 4, p. 412–421, 2014. Disponível em: <https://doi.org/10.1080/01639374.2014.882465>. Acesso em: 25 mar. 2025.

BIZER, C.; HEATH, T.; BERNERS-LEE, Tim. Linked data: the story so far. **International Journal on Semantic Web and Information Systems**, v. 5, n. 3, p. 1-22, 2009. Disponível em: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>. Acesso em: 26 mar. 2025.

BELLOTTO, A.; BEKESI, J. Enriching Metadata for a University Repository by Modelling and Infrastructure: A New Vocabulary Server for Phaidra. **Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare**, v. 72, n. 2, p. 446-459, 2019.

BORKO, H. Information science. What is it? **American Documentation**, v.19, n.1, p.3-5, 1968. Disponível em: <https://www.marilia.unesp.br/Home/Instituicao/Docentes/EdbertoFerneda/mri-01---information-science---what-is-it.pdf>. Acesso em: 26 mar. 2025.

BUCKLAND, M. K. Information as thing. **Journal of the American Society for information science**, v. 42, n. 5, p. 351-360, 1991. Disponível em: [https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/\(SICI\)1097-4571\(199106\)42%3A5%3C351%3A%3AAID-ASI5%3E3.0.CO%3B2-3](https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1097-4571(199106)42%3A5%3C351%3A%3AAID-ASI5%3E3.0.CO%3B2-3). Acesso em: 26 mar. 2025.

BUTLER, J. **Gender trouble**: Feminism and the subversion of identity. New York, NY: Routledge, 1990.

CANDELA, G. *et al.* A linked open data framework to enhance the discoverability and impact of culture heritage. **Journal of Information Science**, v. 45, n. 6, p. 756-766, 2019.

CHEN, P. Entity-relationship modeling: historical events, future trends, and lessons learned. In: **Software pioneers**, 2002. p. 296-310. Disponível em: http://www.csc.lsu.edu/~chen/pdf/Chen_Pioneers.pdf. Acesso em: 26 mar. 2025.

CHEN, S. Semantic enrichment of linked archival materials. **KO KNOWLEDGE ORGANIZATION**, v. 46, n. 7, p. 530-547, 2020.

DI BITONTO, P. *et al.* A recommendation system to promote local cultural heritage. **Journal of e-learning and Knowledge Society**, v. 7, n. 3, p. 97-107, 2011.

DÜRING, M. *et al.* Impresso inspect and compare. comparação visual de artigos de jornais históricos semanticamente enriquecidos. **Information**, v. 12, n. 9, p. 348, 2021.

DÜRING, M. *et al.* impresso Text Reuse at Scale. An interface for the exploration of text reuse data in semantically enriched historical newspapers. **Frontiers in big Data**, v. 6, p. 1249469, 2023.

EUROPEAN COMMISSION. Directorate-General for Communications Networks, Content and Technology. **Women in the digital age**: executive summary.

Publications Office, 2018. Disponível em: <https://data.europa.eu/doi/10.2759/517222>. Acesso em: 26 mar. 2025.

GARRIDO ARILLA, M. R. **Teoría e historia de la catalogación de documentos**. Madrid: Síntesis, 1999.

GEORGIADIS, H. *et al.* Using type and temporal semantic enrichment to boost content discoverability and multilingualism in the Greek cultural aggregator SearchCulture. gr. **International Journal of Metadata, Semantics and Ontologies**, v. 13, n. 1, p. 75-92, 2018.

GIL, A. C. **Métodos e técnicas de pesquisa**. 6 ed. São Paulo: Atlas, 2008. 216 p.

GILLILAND, A. J. Setting the Stage. In: BACA, M. (Org.). **Introd. Metadata**. 3. ed. Los Angeles: Getty Research Institute, 2016.

GORMAN, M. Authority control in the context of bibliographic control in the electronic environment. **Cataloging & Classification Quarterly**, New York, v. 38, n. 3/4, p. 11-22, 2004. DOI: https://doi.org/10.1300/J104v38n03_03. Acesso em: 26 mar. 2025.

GREENBERG, J. A quantitative categorical analysis of metadata elements in image-applicable metadata schemas. **Journal of the American Society for Information Science and Technology**, v. 52, n. 11, p. 917-924, 2001.

GREENBERG, J. Understanding Metadata and Metadata Schemes. **Cataloging & Classification Quarterly**, New York, v. 40, n. 3/4, p. 17-36, 2005. DOI: https://doi.org/10.1300/J104v40n03_02

GUINCHAT, C., MENOUE, M. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. Corrigida e aumentada por Marie-France Blanquet. Trad. Míriam Vieira da Cunha. Brasília: IBICT, 1994.

HERRERO PASCUAL, C. El control de autoridades. **Anales de Documentación**, n. 2, p. 121-136, 1999. Disponível em: <https://revistas.um.es/analesdoc/article/view/2621/2601>. Acesso em: 26 mar. 2025.

HINZE, A. *et al.* Capisco: low-cost concept-based access to digital libraries. **International Journal on Digital Libraries**, v. 20, p. 307-334, 20

HITZLER, P. A review of the semantic web field. **Communications of the ACM**, v. 64, n. 2, p. 76-83, fev. 2021. Disponível em: <https://doi.org/10.1145/3397512>. Acesso em: 30 dez. 2024.

HITZLER, P.; KRÖTZSCH, M.; RUDOLPH, S. **Foundations of Semantic Web technologies**. Boca Raton: CRC Press, 2010.

IFLA. **Declaração dos Princípios Internacionais de Catalogação**. Haia: IFLA, 2016. Disponível em: https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp_2016-pt.pdf. Acesso em: 26 mar. 2025.

IFLA. **Functional Requirements for Authority Data**: a conceptual model: Working Group on Functional Requirements and Numbering of Authority Records (FRANAR) Final Report December 2008: As amended and corrected through July 2013. Den Haag, Netherlands: IFLA, 2013. 54 p.

IFLA. Study Group on the Functional Requirements for Bibliographic Records. **Functional Requirements for Bibliographic Records**: final report. 1998. Disponível em: <https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr/frbr.pdf>. Acesso em: 26 mar. 2025.

IFLA. **Functional requirements for subject authority data (FRSAD)**: a conceptual model. Marcia Lei Zeng et al. (ed.). IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR): München: De Gruyter Saur, 2011. (IFLA Series on Bibliographic Control, v. 43). Disponível em: <https://www.ifla.org/node/5849>. Acesso em: 26 mar. 2025.

IFLA. **UNIMARC Authorities Format Manual**. Permanent UNIMARC Committee (PUC). Den Haag, Netherlands: IFLA, 2024. n.p. Disponível em: <https://repository.ifla.org/server/api/core/bitstreams/3a3df76b-dad8-4a41-95ee-5198813784fa/content>. Acesso em: 06 jul. 2025.

ISAAC, A. *et al.* (2015). **Report on enrichment and evaluation**. The Hague, Netherlands: Europeana Task Force on Enrichment and Evaluation. Disponível em: http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/Enrichment_Evaluation/FinalReport_EnrichmentEvaluation_102015.pdf. Acesso em: 26 mar. 2025.

ISOTANI, S.; BITTENCOURT, I. I. **Dados abertos conectados**. São Paulo: Novatec, 2015. Disponível em: <https://ceweb.br/livros/dados-abertos-conectados>. Acesso em: 26 mar. 2025.

JESUS, A. F. **Recomendações teórico-metodológicas para a publicação de dados bibliográficos abertos e conectados**. 2021. 166 p. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de São Carlos, São Carlos, 2021. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/14228>. Acesso em: 26 mar. 2025.

JESUS, A. F.; CASTRO, F. F.; RAMALHO, R. A. S.. O papel das bibliotecas no linked data. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 26, 2021.

JOUDREY, D. N.; TAYLOR, A. G.; WISSER, K. M. **The organization of information**. Fourth edition ed. Santa Barbara, California: Libraries Unlimited, 2018. (Library and information science text series).

KOLLIA, I. *et al.* A systemic approach for effective semantic access to cultural content. **Semantic Web**, v. 3, n. 1, p. 65-83, 2012.

LIBRARY OF CONGRESS. **MARC 21 format for authority data**. 2018. Disponível em: <https://www.loc.gov/marc/authority/ecadhome.html>. Acesso em: 26 mar. 2025.

LIBRARY OF CONGRESS. **375 (Gender) R**. 2022. <https://www.loc.gov/marc/authority/ad375.html>. Acesso em: 26 mar. 2025.

LIBRARY OF CONGRESS. Network Development and MARC Standards Office. **Formato MARC 21 para dados de autoridade**. Tradução: Ana Rafaela Sales de Araújo, Denise Mancera Salgado, Fabrício Silva Assumpção, Felipe Augusto Arakaki, Luciana Cândida da Silva, Marcelo Votto Texeira, Oscar Eliel, Raildo de Sousa Machado. São Paulo: FEBAB, 2021. Título original: MARC 21 Format for Authority Data. ISBN 978 85 85024 13 0. Disponível em: <http://marc.febab.org/>. Acesso em: 10 jan. 2025.

LIRA, M. A. B. **Uma Abordagem Para Enriquecimento Semântico de Metadados Para Publicação de Dados Abertos**. 2014. 95 f. Dissertação (Mestrado em Ciência da Informação) - Centro de Informática, Universidade Federal de Pernambuco, Recife, 2014. Disponível em: <https://repositorio.ufpe.br/bitstream/123456789/11570/1/DISSERTA%c3%87%c3%83O%20M%c3%a1rcio%20Angelo%20Bezerra%20de%20Lira.pdf>. Acesso em: 26 mar. 2025.

MANGUINHAS, H. *et al.* **Exploring comparative evaluation of semantic enrichment tools for cultural heritage metadata**. In: *Research and advanced technology for digital libraries: 20th International Conference on Theory and Practice of Digital Libraries, TPDL 2016*, Hannover, 5–9 set. 2016. Proceedings 20. Springer International Publishing, 2016. p. 266–278.

MARCONDES, C. H. Representação e economia da informação. **Ci. Inf.**, Brasília, v. 30, n. 1, p. 61-70, jan./abr. 2001. Disponível em: <http://www.scielo.br/pdf/ci/v30n1/a08v30n1>. Acesso em: 26 mar. 2025.

MARTIN, J. M. **Records, responsibility, and power: an overview of cataloging ethics**. *Cataloging & Classification Quarterly*, v. 59, n. 2-3, p. 281–304, 2021. DOI: 10.1080/01639374.2020.1871458. Acesso em: 26 mar. 2025.

MENDONÇA, A. C. N. **O linked data e a Europeana**: uma análise da estrutura de dados de uma coleção de obras de arte de artistas femininas. 2022. 65 f., il. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) – Universidade de Brasília, Brasília, 2022. Disponível em: <https://bdm.unb.br/handle/10483/33154>. Acesso em: 7 jul. 2025.

MILLER, S. J. **Metadata for digital collections**. New York: Neal-Schuman Publishers, 2011. (How-to-do-it manuals, no. 179).

OLSON, H. A. The Power to Name: Representation in Library Catalogs. **Signs**, v. 26, n. 3, p. 639–668, 2001. Disponível em: <http://www.jstor.org/stable/3175535>. Acesso em: 26 mar. 2025.

PCC AD HOC TASK GROUP ON GENDER IN NAME AUTHORITY RECORDS.

Report of the PCC Ad Hoc Task Group on Gender in Name Authority Records.

Library of Congress, 2016. Disponível em:

https://www.loc.gov/aba/pcc/documents/Gender_375%20field_RecommendationReport.pdf. Acesso em: 25 mar. 2025.

PCC AD HOC TASK GROUP ON GENDER IN NAME AUTHORITY RECORDS.

Revised report on recording gender in personal name authority records. Library

of Congress, 2022. Disponível em: [https://www.loc.gov/aba/pcc/documents/gender-](https://www.loc.gov/aba/pcc/documents/gender-in-NARs-revised-report.pdf)

[in-NARs-revised-report.pdf](https://www.loc.gov/aba/pcc/documents/gender-in-NARs-revised-report.pdf). Acesso em: 26 mar. 2025.

PENG, Z. *et al.* Author gender metadata augmentation of hathitrust digital library.

Proceedings of the America Society for Information Science and Technology, v.

51, n. 1, p.1-4, 2014. Disponível em:

<https://doi.org/10.1002/meet.2014.14505101098>. Acesso em: 26 mar. 2025.

PETIYA, S. A Linked Open Data model for describing comic book sequences:

Exploring semantic enrichment opportunities with graphic medicine. **Art Libraries**

Journal, v. 48, n. 3, p. 74-79, 2023.

PFEFFER, J.; SUTTON, R. I. **Hard facts, dangerous half-truths, and total**

nonsense: Profiting from evidence-based management. Harvard Business Press, 2006.

POMERANTZ, Jeffrey. **Metadata**. Massachusetts: The MIT Press, 2015. 239p.

POSSEMATO, T. Recording gender in the person entity: an ongoing discussion.

JLIS.it , [S. l.], v. 16, n. 1, p. 1–12, 2025. DOI: 10.36253/jlis.it-630. Disponível em:

<https://www.jlis.it/index.php/jlis/article/view/630>. Acesso em: 26 mar. 2025.

POUX, F. *et al.* 3D point clouds in archaeology: Advances in acquisition, processing

and knowledge integration applied to quasi-planar objects. **Geosciences**, v. 7, n. 4,

p. 96, 2017.

PRODANOV, C. C.; FREITAS, E. C. de. **Metodologia do trabalho científico**:

métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo:

Feevale, 2013. 276 p. Disponível em: [https://www.feevale.br/institucional/editora-](https://www.feevale.br/institucional/editora-feevale/metodologia-do-trabalhocientifico--2-edicao)

[feevale/metodologia-do-trabalhocientifico--2-edicao](https://www.feevale.br/institucional/editora-feevale/metodologia-do-trabalhocientifico--2-edicao).

READ, J. M. **Cataloguing without tears**: managing knowledge in the information

society. Chandos Publishing: Oxford, 2003.

RDF WORKING GROUP. **Resource Description Framework**. W3C, 2014.

Disponível em: <https://www.w3.org/RDF/>. Acesso em: 27 dez. 2024.

RILEY, J. **Understanding metadata**: what is metadata, and what is it for? Baltimore:

National Information Standards Organization (NISO), 2017. Disponível em:

[https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20M](https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf)

[etadata.pdf](https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf). Acesso em: 26 mar. 2025.

RIVA, P; LE BOEUF, P; ŽUMER, M. **IFLA Library Reference Model: a conceptual model for bibliographic Information**. Netherlands: Internacional Federation of Library Associations and Institutions, 2017. Disponível em: https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august2017_rev201712.pdf. Acesso em: 26 mar. 2025.

ROBINSON, S. et al. **Conceptual modeling for discrete-event simulation**. CRC Press, Inc., 2010.

SANTOS, P. L. V. A. C. Catalogação, formas de representação e construções mentais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 6, n. 1, p. 1-24, 2013. Disponível em: <<http://hdl.handle.net/11449/115044>>. Acesso em: 27 dez. 2024.

SANTOS, P. L. V. A. C; SANT'ANA, R. C. G. Dado e Granularidade na perspectiva da Informação e Tecnologia: uma interpretação pela Ciência da Informação. **Ciência da Informação**, v. 42, n. 2, p. 199–209, 2013. Disponível em: <https://revista.ibict.br/ciinf/article/view/1382>. Acesso em: 25 mar. 2025.

SANTOS, P. L. V. A. C.; SIMIONATO, A. C.; ARAKAKI, F. A. Definição de metadados para recursos informacionais: apresentação da metodologia BEAM. **Informação & Informação**, [S. l.], v. 19, n. 1, p. 146–163, 2014. DOI: 10.5433/1981-8920.2014v19n1p146. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/15251>. Acesso em: 6 jul. 2025.

SARDO, L. Ethics and cataloguing. **JLIS.it**, [S. l.], v. 10, n. 3, p. 1–17, 2019. DOI: 10.4403/jlis.it-12565. Disponível em: <https://www.jlis.it/index.php/jlis/article/view/63>. Acesso em: 25 mar. 2025.

SALGADO, D. M. *et al.* Formato MARC 21 para dados de autoridade: considerações no processo da tradução. **Revista Brasileira de Biblioteconomia e Documentação**, [S. l.], v. 18, n. 3, p. 1–19, 2022. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1914>. Acesso em: 25 mar. 2025.

SANTARÉM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do Brasil frente às iniciativas internacionais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 8, n. 2, 2015.

SCOTT, J. Gênero: uma categoria útil de análise histórica. **Educação e Realidade**, Porto Alegre, v. 20, n. 2, jul./dez., 1995.

SCHÖCH, C. Big? Smart? Clean? Messy? Data in the Humanities. **Journal of Digital Humanities**, v. 2, n. 3, p. 2–13, 2013. Disponível em: <https://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>. Acesso em: 25 mar. 2025.

SEIKEL; M.; STEELE, T. How MARC has changed: the history of the format and its forthcoming relationship to RDA. **Technical Services Quarterly**, v. 28, n. 3, p. 322-334, 2011. DOI: 10.1080/07317131.2011.574519. Acesso em: 25 mar. 2025.

SILVA, R. E. da. **As tecnologias da web semântica no domínio bibliográfico**. 2013. 134 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Marília, 2013. Disponível em: <http://repositorio.unesp.br/handle/11449/93653>. Acesso em: 25 mar. 2025.

SILVA, M. R.; FUJITA, M. S. L. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, v. 16, p. 133-161, 2004.

SILVA, D. L. da. **Ontologias para representação de documentos multimídia: análise e modelagem**. 2014. 442 f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

SMITH-YOSHIMURA, K. **The rise of Wikidata as a linked data source**. *Hanging Together*, 2018. Disponível em: <https://hangingtogether.org/the-rise-of-wikidata-as-a-linked-data-source/>. Acesso em: 26 mar. 2025.

TAYLOR, A. G.; JOUDREY, D. N. Metadata: access and authority control. **The organization of information**. 3. ed. Westport, Conn.: Libraries Unlimited, 2009.

TILLET, B. B. Authority control: State of the art and new perspectives. **Cataloging & Classification Quarterly**, v. 38, n. 3-4, p. 23-41, 2004b. doi:10.1300/J104v38n03_04. Disponível em: https://www.tandfonline.com/doi/abs/10.1300/J104v38n03_04. Acesso em: 26 mar. 2025.

TILLET, B. B. Keeping Libraries Relevant in the Semantic Web with RDA: Resource Description and Access. **JLIS.It**, v.2, n. 2, 2011. Doi.org/10.4403/jlis.it 4760. Disponível em: <https://jlis.fupress.net/index.php/jlis/article/view/302>. Acesso em: 26 mar. 2025.

THOMPSON, K. J. More than a name: a content analysis of name authority records for authors who self-identify as trans. **Library Resources & Technical Services**, Chicago, v. 60, n. 3, p. 140-155, July 2016. Disponível em: <https://journals.ala.org/index.php/lrts/article/view/6036/7732>. Acesso em: 2 jul. 2025.

TRANS METADATA COLLECTIVE *et al.* Metadata best practices for trans and gender diverse resources. **Librarian Publications & Presentations**, n. 52, 2022. Disponível em: <https://digitalcommons.osgoode.yorku.ca/librarians/52>. Acesso em: 9 jun. 2025.

UNESCO. Equals Global Partnership. **I'd blush if I could**: closing gender divides in digital skills through education. UNESCO; EQUALS, 2019. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1%C2%A0>. Acesso em: 26 mar. 2025.

VAN VEEN, T.; LONIJ, J.; KOPPELAAR, H.. Semantic enrichment: a low-barrier infrastructure and proposal for alignment. **D-Lib Magazine**, v. 21, n. 7/8, 2015.

VARTHIS, E. *et al.* Automatic metadata extraction via image processing using Migne's Patrologia Graeca. **International Journal of Metadata, Semantics and Ontologies**, v. 14, n. 4, p. 265-278, 2020.

VARTHIS, E. *et al.* Semantic enrichment on large scanned collections through their "satellite texts": the paradigm of Migne's Patrologia Graeca. **Information Discovery and Delivery**, v. 50, n. 2, p. 217-234, 2022.

VLACHIDIS, A. *et al.* Automatic metadata generation in an archaeological digital library: Semantic annotation of grey literature. **Computational Linguistics: Applications**, p. 187-202, 2013.

OLIVER, C. **Introducing RDA: a guide to the basics after 3R**. Second edition ed. Chicago: ALA Editions, 2021.

OLSON, H. A. The Power to Name: Representation in Library Catalogs. **Signs**, v. 26, n. 3, p. 639–668, 2001. Disponível em: <http://www.jstor.org/stable/3175535>. Acesso em: 26 mar. 2025.

ONU. **Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável**. 2015. Disponível em: <https://brasil.un.org/sites/default/files/2020-09/agenda2030-pt-br.pdf>. Acesso em: 26 mar. 2025.

WANG, X. *et al.* Data modeling and evaluation of deep semantic annotation for cultural heritage images. **Journal of Documentation**, v. 77, n. 4, p. 906-925, 2021.

WOOD, D.; ZAIDMAN, M.; RUTH, L. **Linked Data: structured data on the Web**. EUA: Manning Publications, 2014.

WORLD WIDE WEB CONSORTIUM. **Best practices for publishing Linked Data**. 2014. Disponível em: <https://www.w3.org/TR/ld-bp>. Acesso em: 26 mar. 2025.

WORLD WIDE WEB CONSORTIUM. **Data on the web best practices**. 2019. Disponível em: <https://www.w3.org/Translations/DWBP-pt-BR/#dataVocabularies>. 2019. Acesso em: 26 mar. 2025.

WORLD WIDE WEB CONSORTIUM. **SPARQL**. 2013. Disponível em: <https://www.w3.org/TR/sparql11-query/>. Acesso em: 26 mar. 2025.

ZENG, M. L. Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article. **Profesional de la información, [S. l.]**, v. 28, n. 1, 2019. DOI: 10.3145/epi.2019.ene.03. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2019.ene.03>. Acesso em: 26 mar. 2025.

ZENG, M.L.; QIN, J. **Metadata**. New York: Neal-Schuman Publishers, 2008. 365 p.

ZENG, M.L; QIN, J. **Metadata**. 2.ed. New York: Neal-Schuman Publishers, 2016. 584 p.