



**UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

JEFFERSON LEITE OLIVEIRA FERREIRA

**DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS EM MARC 21:
ESTRATÉGIAS DE CONVERSÃO PARA FORMATOS DE
INTERCÂMBIO MULTIPLATAFORMA**

**FORTALEZA
2018**

JEFFERSON LEITE OLIVEIRA FERREIRA

DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS EM MARC 21:
ESTRATÉGIAS DE CONVERSÃO PARA FORMATOS DE INTERCÂMBIO
MULTIPLATAFORMA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Representação e Mediação da Informação e do Conhecimento.

Linha de Pesquisa: Representação da Informação e Conhecimento e Tecnologias

Orientador: Profa. Dra. Virgínia Bentes Pinto.

FORTALEZA
2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

F441d

Ferreira, Jefferson Leite Oliveira.

Descrição de recursos bibliográficos em MARC 21 : estratégias de conversão para formatos de intercâmbio multiplataforma / Jefferson Leite Oliveira Ferreira. – 2018.

101 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Ciência da Informação, Fortaleza, 2017.

Orientação: Profª. Dra. Virgínia Bentes Pinto.

1. Representação Descritiva. 2. Metadados Descritivos. 3. Conversão de Metadados. 4. Cabeçalhos de Autoridade.

I. Título.

CDD 025

JEFFERSON LEITE OLIVEIRA FERREIRA

DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS EM MARC 21:
ESTRATÉGIAS DE CONVERSÃO PARA FORMATOS DE INTERCÂMBIO
MULTIPLATAFORMA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Ceará, como requisito parcial à qualificação do mestrado.

Área de concentração: Representação e Mediação da Informação e do Conhecimento.

Linha de Pesquisa: Representação da Informação e Conhecimento e Tecnologias

Orientador: Profa. Dra. Virgínia Bentes Pinto.

Aprovada em: 17/08/2018.

BANCA EXAMINADORA

Profa. Dra. Virgínia Bentes Pinto (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dra. Zaira Regina Zafalon
Universidade Federal de São Carlos (UFSCar)

Prof. Dr. Heliomar Cavati Sobrinho
Departamento de Ciências da Informação (DCINF- UFC)

Profa. Dra. Andréa Soares Rocha da Silva
Departamento de Ciências da Informação (DCINF- UFC)

Prof. Dr. Guilherme Ataíde Dias
Universidade Federal da Paraíba (UFPB)

AGRADECIMENTOS

Começo meus agradecimentos com a justa e necessária menção à Universidade Federal do Ceará, por meio do Programa de Pós-Graduação em Ciência da Informação (PPGCI), tornou esta valorosa jornada de pesquisa possível.

Agradeço à Profa. Dra. Virgínia Bentes Pinto não apenas pela orientação, sempre eivada de conhecimentos e prestimosas reflexões, mas por sua firme atuação como pesquisadora, cuja presença se confunde com minha trajetória acadêmica e se consolida sempre como referência a ser seguida.

A todos os alunos, professores e demais atores que colaboram direta ou indiretamente com o PPGCI cientes de que sua parcela de contribuição em muito enaltece e contribui com a consolidação deste tão esperado Mestrado Acadêmico.

Aos estimados colegas: Ana Pricila Celedônio, Bárbara Luísa Ferreira Carneiro, Felipe Alves de Lima Braga, Giordana Nascimento de Freitas e Silva, Igor Peixoto Torres Girão, Ismael Lopes Mendonça, Joana D'arc Páscoa Bezerra Fernandes, Morgana Ramos Albuquerque e Natanna Santana de Moraes que me legaram o privilégio de suas presenças e com quem compartilhei o orgulho de sermos a primeira turma deste valoroso mestrado acadêmico.

Reconheço aqui na pessoa da Ana Elizabeth Albuquerque Maia, diretora da Biblioteca de Ciências Humanas da Universidade Federal do Ceará, uma profissional que jamais se furtou a motivar e acompanhar a qualificação dos quadros de servidores que integram esta Grande Biblioteca da qual me orgulho a cada dia em fazer parte.

Agradeço a Veruska Maciel, secretária do PPGCI, pelos suportes de natureza administrativa em todas as ocasiões em que pairavam dúvidas sobre procedimentos internos ao Programa.

Meus cordiais cumprimentos aos bibliotecários Francisco Edvander Pires Santos e Ericson Bezerra Viana pelas assistências no envio de documentos que se fizeram necessários no decorrer desta caminhada.

RESUMO

A pesquisa objetiva pesquisar estratégias de conversão de registros bibliográficos do formato MARC 21 para os esquemas de metadados MARCXML, *Metadata Authority Description Schema*(MADS) e *Metadata Object Description Schema*(MODS), por meio dos cabeçalhos de autoridade de nomes e assuntos como categorias de análise. Procedeu-se ao mapeamento dos subcampos com seus respectivos valores e o posterior cotejamento com as *tags* dos esquemas de metadados alvo da conversão. A pesquisa de caráter exploratória e descritiva propiciada por levantamentos bibliográficos e documentais adota a observação sistemática como técnica viabilizadora da etapa empírica. Aborda o modelo entidade-relacionamento adotado pelos *Functional Requirements for Bibliographic Records* (FRBR) como ferramenta para modelar o relacionamento entre os atributos dos metadados em estudo e o formato MARC com as prescrições dos *Functional Requirements for Authority Data* (FRAD) para os dados de autoridade de nome pessoal e os *Functional Requirements for Subject Authority Data* (FRSAD) para os dados de autoridade de assuntos tópicos. A análise preliminar demonstrou que a flexibilidade dos esquemas permite ampliar o escopo descritivo dos catálogos de autoridade, bem como viabiliza a abordagem por meio de remissivas e qualificadores apropriados com ênfase nos aspectos de caráter relacional. Analisa a flexibilidade notacional de esquemas correlatos como o *Metadata Encoding and Transmission Standard* (METS) na esteira da perspectiva de recursos de informação interoperáveis. Insere o vocabulário *Dublin Core* na discussão e codificação de metadados descritivos, o que possibilitou mapear limites e possibilidades deste vocabulário no contexto bibliográfico. Os resultados das conversões realizadas ilustraram o papel estratégico do modelo entidade-relacionamento como paradigma conceitual adotado pelos FRBR e, subsidiariamente pelos FRAD e FRSAD, na condução das etapas de mapeamento e codificação dos dados de autoridade de nome pessoal e assunto tópico de MARC para MADS e MODS. Foi observado que todos os campos e subcampos MARC analisados, além de encontrarem seus equivalentes nos esquemas XML correspondentes, são acrescidos de atributos relacionais que demonstraram efetividade nos vínculos que associam autores a papéis assumidos nos recursos. Conclui que as estratégias de conversão aplicadas extraem o valor semântico do paradigma relacional preconizado pelos FRBR como elemento chave para alcance de expressividade, coerência interna e sofisticação das relações sintáticas e semânticas tão caras à efetividade dos catálogos. Sugere, a partir da literatura consultada e dos resultados verificados na etapa empírica, pesquisas que visem contemplar estratégias de conversão de metadados com ênfase nos grupos 2 e 3 dos FRBR.

Palavras-chave: Representação Descritiva. Metadados Descritivos. Conversão de Metadados. Cabeçalhos de Autoridade.

ABSTRACT

The objective of this research is to define strategies for converting bibliographic records from MARC 21 format to MARCXML, *Metadata Authority Description Schema* (MADS) and *Metadata Object Description Schema* (MODS) metadata schemas, through name and subject authority as analysis categories. The mapping of the subfields with their respective values and the subsequent collation with the tags of the metadata schemes target of the conversion were done. The exploratory and descriptive research provided by bibliographical and documentary surveys adopts systematic observation as a technique that enables the empirical stage. It addresses the entity-relationship model adopted by the *Functional Requirements for Bibliographic Records* (FRBR) as a tool to model the relationship between the attributes of the metadata under study and the MARC format with the requirements of *Functional Requirements for Authority Description* (FRAD) and the *Functional Requirements for Subject Authority Data* (FRSAD) for the topics subject authority data. The preliminary analysis showed that the flexibility of the schemas allows to treat redundancies and ambivalences in the representation of personal names and their attributes. It also allows the approach through appropriate references and qualifiers with emphasis on aspects related to synonymy and polysemy of subject headings. It analyzes the notational flexibility of related schemas such as *Metadata Encoding and Transmission Standard* (METS) in the wake of the perspective of interoperable information resources. It inserts the *Dublin Core* vocabulary into the discussion and coding of descriptive metadata, which made it possible to map the limits and possibilities of this vocabulary in the bibliographic context. The results of the conversions carried out illustrated the strategic role of the entity-relationship model as a conceptual paradigm adopted by the FRBR and, in a subsidiary way, by the FRAD and FRSAD, in the process of mapping and coding the personal name and subject matter authority data from MARC to MADS and MODS. It was observed that all the analyzed MARC fields and subfields, besides finding their equivalents in the corresponding XML schemas, are added with relational attributes that demonstrated effectiveness in the links that associate authors to roles assumed in the resources. It concludes that the applied conversion strategies extract the semantic value of the relational paradigm advocated by the FRBR as a key element for the reach of expressiveness, internal coherence and sophistication of the syntactic and semantic relations so dear to the effectiveness of the catalogs. It suggests, from the literature consulted and the results verified in the empirical stage, researches that aim to contemplate strategies of metadata conversion with emphasis in groups 2 and 3 of FRBR.

Keywords: Descriptive Representation. Descriptive Metadata. Metadata Conversion. Subject Authority Headers. Personal Names Authority Headers.

LISTA DE ILUSTRAÇÕES

Figura 1 - Representação de indicadores e sinalizadores em campo de autoridade de nome pessoal.....	31
Figura 2 - Sintaxe de um documento XML.....	33
Figura 3 - Sintaxe de um DTD interno.....	35
Figura 4 - Sintaxe de um DTD externo.....	35
Figura 5 - Sintaxe XML Schema.....	37
Figura 6 - Sintaxe do esquema METS.....	39
Figura 7 - Posição do MARCXML na escala de conversão de metadados.....	41
Figura 8 - Áreas de título e assunto tópico em MARCXML.....	42
Figura 9 - Relacionamento muitos para muitos entre <i>work ethema</i>	53
Figura 10 - Relacionamentos possíveis entre <i>work, thema e nomen</i>	54
Figura 11- Hierarquia de elementos de autoridade de nome em MADS.....	56
Figura 12 - Fundamentos para o Modelo Conceitual.....	58
Figura 13 - Mapeamento de dados atribuídos a Autor Pessoal.....	59
Figura 14 - Mapeamento de um registro de autoridade de nome pessoal.....	61
Figura 15 - Termo Tópico e remissivas aplicáveis a subdivisões.....	63
Figura 16 - Aplicação do MARCXML na codificação de dados de autoridade de nome pessoal e assunto tópico.....	67
Figura 17 -Aplicação do MARCXML na codificação de dados de autoridade de nome pessoal e remissivas correspondentes.....	68
Figura 18 - Esquema MADS baseado em autoridade de nome pessoal.....	70
Figura 19 - Aplicação do MADS na codificação de dados de autoridade de nome pessoal.....	71
Figura 20 - Interface de busca no catálogo de autoridade da <i>Library of Congress</i>	72
Figura 21 - Codificação de subelemento da autoridade de nome.....	73
Figura 22 - Aplicação do MADS na codificação dos atributos relacionados à nome pessoal.....	76
Figura 23 - Esquema MODS de registro consultado no catálogo da LOC.....	82
Figura 24 - Aplicação do MODS na codificação de identificadores como remissivas.....	85
Figura 25 - Conversão para subelementos de <i>subject</i> em MODS.....	86

LISTA DE ABREVIATURAS E SIGLAS

AACR2	Anglo-American Cataloguing Rules
CSS	Cascading Style Sheets
DC	Dublin Core
DTD	Document Type Definition
FRBR	Funcional Requirements for Bibliographic Records
HTML	HyperText Markup Language
IFLA	International Federation of Library Associations and Institutions
ISBD	International Standard Bibliographic Description
ISO2709	Documentation Format for Bibliographic Interchange on Magnetic Tape
ISO8879	Information processing - Text and office systems
ISAN	International Standard Author Number
JSC	Joint Steering Committee
LCSH	Library of Congress Subject Headings
LOM	Learning Object Metadata
MARC	Machine Readable Cataloguing
MADS	Metadata Authority Description Schema
METS	Metadata Encoding Transmission Standard
MODS	Metadata Object Description Schema
NACO	Name Authority Cooperative Program
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OPAC	Online Public Access Catalogue
RDA	Resource Description and Access
RDF	Resource Description Framework
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
SGML	Standard Generalized Markup Language
SKOS	Simple Knowledge Organization System
TDIC	Tecnologias Digitais de Informação e de Comunicação
TCP/IP	Transmission Control Protocol/Internet Protocol
UMLS	Unified Medical Language System
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
UTF-8	8-bit Unicode Transformation Format
VRA Core	Visual Resources Association Core Categories for Visual Resources
W3C	World Wide Web Consortium
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language for Transformation

LISTA DE QUADROS

Quadro 1 - Delimitação de campos para representação de autoridades de nome pessoal.....	16
Quadro 2 – Forma da entrada autorizada de nome e assunto segundo indicadores.....	31
Quadro 3 - Atributos e subelementos dos principais <i>namespaces</i> do esquema MADS.....	43
Quadro 4 - Relacionamentos entre nomes pessoais, de família, entidade e obra.....	49
Quadro 5 - Atributos de autoridade de nome pessoal.....	50
Quadro 6 - Tarefas gerais do usuário no escopo do FRBR.....	51
Quadro 7 - Mapeamento dos atributos de nome segundo FRAD.....	57
Quadro 8 - Registro MARC para uma coletânea.....	60
Quadro 9 - Relação de campos e subcampos para definição de hierarquia de assuntos.....	65
Quadro 10 - Descrição dos <i>namespaces</i> em MADS e respectivas equivalências MARC.....	75
Quadro 11 - Conversão de campos de assunto MARC para MODS.....	78
Quadro 12 - Descrição dos <i>namespaces</i> em MODS e respectivas equivalências em MARC.....	80
Quadro 13 - Subelementos MODS expressos em <i>Dublin Core</i>	81
Quadro 14 - Subelementos MODS aplicados à descrição de assunto.....	83

SUMÁRIO

1 INTRODUÇÃO	13
1.1 Problema da pesquisa	15
1.2 Justificativa	18
1.3 Objetivo Geral	21
1.4 Objetivos específicos	21
1.5 Estrutura da dissertação	21
2 DIRETRIZES METODOLÓGICAS: MÉTODOS, TÉCNICAS E O UNIVERSO DA PESQUISA	23
3 SOBRE METADADOS DESCRITIVOS	26
3.1 Origem	26
3.1.1 Aplicabilidade	28
<i>3.1.1.1 Formato MARC</i>	29
<i>3.1.1.2 Linguagem XML</i>	32
3.2 Especificações de codificação e controle	34
3.2.1 Document Type Definition (DTD)	34
3.2.2 XML Schema	36
4 ESQUEMAS DE METADADOS APLICADOS À DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS	38
4.1 Metadata Encodings and Transmission Standard (METS)	38
4.2 MARCXML	40
4.3 Metadata Authority Description Schema (MADS)	42
4.4 Metadata Object Description Schema (MODS)	44
5 MODELAGEM CONCEITUAL DE METADADOS PARA O DOMÍNIO BIBLIOGRÁFICO	45
5.1 Functional Requirements for Bibliographic Records (FRBR)	46
5.2 Functional Requirements for Authority Records (FRAD)	48
5.3 Functional Requirements for Subject Authority Data (FRSAD)	51
6 ESTUDO EMPÍRICO: MAPEAMENTO E CONVERSÃO DOS PONTOS DE ACESSO DE AUTORIDADES DE NOME E ASSUNTO	55
6.1 Mapeamento das autoridades de nomes	55
6.2 Mapeamento das autoridades de assuntos	62
6.3 Codificação em MARCXML	66

6.4 Codificação em MADS.....	69
6.5 Codificação em MODS.....	78
7 ANÁLISE DOS DADOS E DISCUSSÃO DOS RESULTADOS.....	88
8 CONCLUSÕES.....	91
REFERÊNCIAS.....	93

1 INTRODUÇÃO

A catalogação tem sua origem entre os anos de 668 e 626 a.C na forma de uma escrita a seu respeito na coleção da Biblioteca de Assurbanipal no reino da Assíria. Mas foi na segunda metade do século XX e nas primeiras décadas deste século que se registrou avanço significativo na prática da catalogação, sobretudo em resposta ao crescimento exponencial da informação registrada em paralelo à diversidade dos suportes informacionais, o que conduziu e conduz a necessidade de revisão constante das regras para representação descritiva de recursos de informação. Para o campo da Biblioteconomia a catalogação assume um papel imprescindível, uma vez que trata de estabelecer regras formais de descrição de uma miríade de tipos documentais, para que estes se tornem passíveis de serem representados e diferenciados dentro de um universo documental. À catalogação também cabe definir dados de localização e pontos de acesso por meio dos quais estas fontes possam ser recuperadas. (MACHADO, 2013, p. 3)

Foi em 1961 com a declaração dos **Princípios de Paris** que ficaram estabelecidas as regras formais da catalogação. No que diz respeito às funções do catálogo constantes na declaração, Silveira e Tálamo (2009, p. 112) lembram que o catálogo deve proporcionar ao usuário a localização de um item, seja por autor, título ou seu substituto conveniente.

Com o aperfeiçoamento das diretrizes da declaração, conduzidas por especialistas no âmbito da *International Federation of Library Associations and Institutions* (IFLA) em seu relatório datado de 1997, foi recomendada a adoção do padrão *International Standard Bibliographic Description* (ISBD) como esforço cooperativo internacional que viesse a contemplar todos os países, além de estender seu escopo para museus, arquivos e outras instituições. A premissa fundamental do ISBD enquanto acordo de nível internacional consiste, além da efetivação do controle bibliográfico a cargo das agências nacionais, em prover elementos descritivos mínimos necessários para identificar e selecionar recursos de informação. Possui aplicabilidade direta nas regras de pontuação e descrição distribuídas nas oito áreas que integram o *Anglo-American Cataloguing Rules* (AACR2), dentre elas o título e a indicação de responsabilidade. Com isto, a interoperabilidade passa a ser outro grande objetivo destacado pela IFLA baseado no ISBD, a exemplo de regras uniformes para os pontos de acesso da descrição.

Com a revisão das regras de catalogação a cargo do *Joint Steering Committee* (JSC), que segundo Gómez (2012, p. 2) ao lançar em 2004 a terceira edição do AACR2 e de ter recebido críticas da comunidade de catalogadores, sentiu-se a seguir a necessidade de criar regras mais consistentes capazes de representar com maior acurácia a coerência interna do domínio bibliográfico. Essas regras ficaram a cargo da IFLA que estabeleceu requisitos funcionais para registros bibliográficos, ou FRBR.

A partir do vertiginoso aumento do volume de registros bibliográficos nos acervos das bibliotecas, iniciativas para tratar adequadamente a redundância dos cabeçalhos de autoridade como os de nomes, ensejaram a criação de modelos conceituais e programas cooperativos a serem adotados internacionalmente. Neste contexto, Snyman e Van Rensburg (1999, p.314) abordam o *International Standard Author Number* (ISAN) e suas potencialidades para agregar e relacionar autores a atributos de nacionalidade, idioma, assuntos e publicações. A IFLA, através dos Requisitos Funcionais para Dados de Autoridade ou FRAD, destaca do universo bibliográfico os atributos que representam o autor, tais como data de nascimento e morte, origem, afiliação, idioma dentre outros. Importa destacar a referência feita por Moulaison (2015, p. 14) ao trabalho desenvolvido pela Biblioteca do Congresso Americano denominado *Name Authority Cooperative Program* (NACO), destinado a prover acordos cooperativos e treinamentos para validar o controle de autoridades de nomes, o que também inclui títulos uniformes de séries e coleções.

A ênfase no conteúdo e na sua descrição impulsionou o delineamento de dois cenários, a saber: a revisão das regras tradicionais da catalogação, definidas pelo AACR2, para um modelo capaz de contemplar a multiplicidade de formatos e suportes de informação, sob a designação de *Resource Description and Access* (RDA) e a modelagem conceitual do universo bibliográfico, apoiado pelo paradigma entidade-relacionamento importado da Ciência da Computação pelos FRBR, sob os auspícios da IFLA. O ponto de ruptura do RDA com o ISBD, tratado anteriormente, e o próprio formato MARC21, é que enquanto estes últimos tratam do armazenamento e/ou apresentação dos dados, aquele tem a função de normalizar conteúdo dos registros de informação independente dos suportes que o contém e de sua forma de visualização. (GÓMEZ, 2012, p. 8).

Em paralelo, os esquemas de metadados digitais aplicados à descrição de registros bibliográficos sofreram profundas e significativas transformações. Desde a estrutura sintática representativa dos pontos de acesso e dos dados de autoridade, como o formato legível por máquina ou MARC21, até a funcionalidade no que concerne ao

intercâmbio e acesso à informação em distintas plataformas, a exemplo dos sistemas de automação de bibliotecas, portais de periódicos e os repositórios institucionais.

Esquemas de metadados baseados na *Extensible Markup Language* (XML), como o *Dublin Core* (DC) e o *Resource Description Framework* (RDF), prestam-se a este propósito, e tem sido amplamente abordados como fator de enriquecimento da representação da informação e de consistência na organização dos índices de entradas de nomes e assuntos.

Esta pesquisa nasce da perspectiva de integração dos metadados digitais no contexto da *World Wide Web*, popularmente conhecida por *web*, ao domínio bibliográfico, com vistas a estabelecer estratégias dinâmicas e flexíveis de representação dos dados de autoridade e assunto na ambiência multiplataforma.

1.1 Problema da pesquisa

Por definição, o formato de intercâmbio de registro bibliográfico MARC foi desenhado para suportar sua legibilidade por máquina como importante requisito para a interoperabilidade entre sistemas. Para sua consecução é aplicada a norma ISO 2709, que nas palavras de Schmidt e Patel (1999, p. 287), é o padrão para representação eletrônica dos dados codificados em MARC.

O padrão ISO 2709, conforme Assumpção (2013, p. 11), atua na codificação dos registros MARC desde sua criação nos anos 60. Sua longevidade e aplicabilidade impulsionaram o desenvolvimento de testes e simulações com formatos semelhantes e esquemas de metadados, bem como a sofisticação da abordagem de formas variantes de nomes e assuntos.

É amplamente reconhecido que a codificação dos campos e subcampos do formato e sua aplicabilidade descritiva, tem servido aos propósitos a que se destina, motivo pelo qual, até os dias atuais, seu uso é mandatório para a conformidade normativa dos sistemas de automação de bibliotecas.

No entanto, são inúmeros os relatos de pesquisa e estudos de caso encontrados na literatura que apontam e descrevem as limitações do formato afetado pela elevada produção do conhecimento e a conseqüente complexidade delegada aos pontos de acesso da descrição, sejam eles para representar dados de autoridade de nomes ou de assunto:

No lado estrutural pode-se concluir que as implementações da ISO 2709 permitem que um extensivo número de diferentes elementos de dados seja organizado em uma simples estrutura de dois níveis, mas o formato em si

não suporta diretamente estruturas mais complexas. Técnicas como links de campos no MARC21 e vínculos embutidos do UNIMARC são meios indiretos para adicionar estruturas mais complexas no MARC, mas está baseado na estrutura expressa através de valores de subcampos. (AALBERG; ZUMER, 2012, p. 853)

Os limites da capacidade descritiva de esquemas de metadados são delimitados pelo contexto de uso. Sobre este ponto, Yee (2004, p. 166) considera que os esquemas representam um modelo de estrutura de dados provenientes de um padrão para sua rotulação e, como tal, serve para isolamento de tipos particulares aplicados a propósitos específicos como indexação ou exibição. Isso implica que mesmo exaustivo na representação dos pontos de acesso da descrição, o formato MARC serve ao propósito da catalogação de tipos particulares de materiais, bem como na delimitação dos campos relativos a dados de autoridade, dentre os quais os de nome e assunto tratados nesta pesquisa. Um exemplo desta delimitação é apresentado por Machado e Brigidi (2013, p. 6) relativo à autoridade de nomes:

Quadro 1 - Delimitação de campos para representação de autoridades de nome pessoal.

040	Fonte da catalogação
100	Nome pessoal
400	Remissiva
VER (se houver)	
670	Fonte positiva de pesquisa
675	Fonte negativa de pesquisa.

Fonte: adaptado de Machado e Brigidi (2013, p. 6)

A flexibilidade na seleção dos campos da descrição do formato MARC para autoridade de nomes, aliada a exaustividade na cobertura nas áreas aplicadas à representação de recursos bibliográficos, ainda contrasta com ausência de regras lógico-relacionais consistentes ao nível da codificação nos registros destinadas à identificação unívoca de autoridades de nomes e assuntos. Tais limitações são analisadas por Tosaka e Park (2013, p. 654) no que concerne a estrutura do formato MARC e sua dificuldade para integração ao modelo entidade-relacionamento do RDA. Neste aspecto, as autoras destacam o advento do *Dublin Core* (DC) e sua abordagem integrada ao RDA, que por meio do RDF, propicia a representação de recursos de

informação de forma consistente em ambientes interoperáveis. Dentre as diversas pesquisas que abordam este aspecto, Machado et. al. (2013, p. 13) e Carrasco, Serrano e Castillo-Buergo (2016, p. 753), após levantamento em bases de autoridades, relataram a presença de muitas entradas de nome (pessoal, entidade e evento) no campo de assunto 650 e a existência de mais de uma entrada de nome para o mesmo autor/entidade/evento por não ser feita pesquisa de nome certo para a fixação da forma da entrada a ser usada no registro incluído na base.

Em seu estudo, Lam(2002, p. 88) apresenta e discute as limitações do formato MARC no tratamento de nomes não latinos que envolvem, além das redundâncias apontadas nas pesquisas acima, a carência de aspectos como identificação de esquemas de romanização aplicados a um campo e a determinação de qual esquema de codificação de caractere deve ser usado em um registro.

Em que concerne ao controle de autoridade, trata-se de uma das etapas mais relevantes para o alcance da consistência na representação de registros bibliográficos. Neste ponto, Assumpção e Santos (2013, p. 109) abordam a diversidade de pontos de acesso para uma entidade, ao considerar que um usuário do catálogo com o conhecimento de somente um dos nomes, pode não encontrar todos os recursos informacionais associados à entidade por ele desejada, a exemplo de todos os livros de um determinado autor. A ausência ou parcialidade na uniformização das entradas autorizadas, são discutidas por Moreno e Brascher (2007, p. 15) que apontam o esforço descentralizado que se caracteriza pela criação de formatos de padronização em cada país, o que vem a dificultar e até inviabilizar o intercâmbio. A mesma observação é feita por Coyle (2010, p. 26), ao ressaltar a tendência contemporânea de diferentes comunidades criarem diferentes metadados para necessidades similares ou idênticas.

Neste particular, resta pouca ou nenhuma chance de reutilização do esquema criado fora da estrutura para a qual fora destinada. Ainda neste aspecto, Aalberg e Zumer (2012, p. 852) apontam as dificuldades que a adoção de novos modelos, ou formatos descentralizados, acarreta na conversão de registros em grandes bases de dados para habilitar a interoperabilidade com novos registros criados. Os autores citam que os dados emergentes constituem grande desafio para uma ampla implementação do modelo conceitual FRBR.

A consistência nas formas de representação da informação do domínio bibliográfico tem encontrado forte amparo nos modelos conceituais do FRBR que, dentre outras finalidades, visam estender a visibilidade dos catálogos em amplas

variedades de suportes e formatos. A partir dessas reflexões, somos conduzidos à questão que norteia a pesquisa. **De que modo as estratégias de conversão de metadados digitais como o MARCXML, *Metadata Authority Description Schema* (MADS) e *Metadata Object Description Schema* (MODS) ao incorporarem atributos da descrição de autoridades de nomes e assuntos do formato MARC, potencializam o enriquecimento lógico-semântico a partir da estrutura conceitual dos FRBR?** Ao tentar responder a esta questão, nos utilizamos dos paradigmas da modelagem conceitual, e em particular, do modelo entidade-relacionamento de Chen (1976) incorporado aos requisitos do FRBR, bem como atrelado às estruturas de metadados para descrição semântica como suporte ao estabelecimento de estratégias para conversão de registros MARC para o ambiente multiplataforma.

1.2 Justificativa

O intercâmbio entre serviços e sistemas biblioteconômicos evoluiu nas últimas décadas das convenções sintáticas e notacionais, para um contexto semântico e contextual. Tal fato foi impulsionado pelo advento da web 2.0, ou web semântica, e as tecnologias aplicadas à descrição e representação de recursos de informação, com particular ênfase, na interoperabilidade e em suas distintas modalidades de manifestação. Como fatos geradores deste paradigma emergente para a Ciência da Informação, surgem novos desafios para o Tratamento Descritivo da Informação (TDI) a despeito de se tratar de tema recorrente na literatura ao longo das últimas décadas:

Apesar de estarmos vivenciando evoluções tecnológicas, é preciso lembrar que a essência do tratamento e recuperação de informações continua a mesma. Continuamos por meio de normas, regras e padrões, buscando maneiras para realizar uma representação e uma organização da informação de modo eficiente. O que tem mudado, no entanto, é o meio tecnológico onde a informação está inserida e necessariamente a alteração nas formas de organização e de representação. (ALVES, 2005, p. 80)

A padronização das entradas autorizadas de nomes e assuntos nos catálogos das bibliotecas, na forma de metadados descritivos, não se encerra apenas numa linguagem de marcação legível por homem e máquina. O que limitaria os problemas da representação dos registros bibliográficos aos limites das convenções sintáticas. Existem, também, desafios de ordem semântica que afetam de forma recorrente a consistência do arranjo dos pontos de acesso:

Nesse sentido, e observando as questões de interoperabilidade semântica entre ambientes informacionais digitais, destaca-se que dentre as heterogeneidades, a que representa um obstáculo maior para a integração de

esquemas e, conseqüentemente, para a integração de dados, é a heterogeneidade semântica, uma vez que esta é marcada por conflitos semânticos que podem ocorrer quando dois contextos não usam a mesma interpretação da informação. (CASTRO; SIMIONATO; ZAFALON, 2016, p. 4)

O cenário multiplataforma, por sua vez, tem o condão de englobar sistemas de automação de bibliotecas, portais educativos e de periódicos e aplicativos móveis. Toda essa infraestrutura informacional encontra nos metadados um requisito indispensável, que em termos de sua aplicabilidade funcional listadas pela *National Information Standards Organization* (NISO) organização para padronização informacional, aponta como principais variáveis:

- Permitir que os recursos sejam encontrados a partir de critérios relevantes;
- Identificação de recursos;
- Agrupar recursos por similaridade;
- Distinguir os itens dessemelhantes;
- Informar a localização. (UNDERSTANDING..., 2004)

Formatos como DC se prestam à catalogação de objetos digitais, o que envolve ampla categoria de materiais. Na sua pesquisa sobre descrição de recursos fotográficos, Gonzales (2014, p. 2), ao justificar a ubiquidade do formato DC, ressalta esta característica no seu projeto destinado a estabelecer links entre as coleções de imagens e os campos do formato MARC correspondentes.

Por sua vez, iniciativas que abordam a linguagem XML para representar coleções e acervos, tem sido objeto de diversas pesquisas que envolvem análise e conversão dos pontos de acesso da descrição e as entradas de autoridade de nomes e assuntos. Em sua pesquisa, Vidakovic e Rackovic (2006, p. 514) apresentam um *framework* desenvolvido na linguagem de programação JAVA destinado a conversão de dados bibliográficos de fichas catalográficas no formato XML, além de discutirem programas de conversão para XML desenvolvidos na Biblioteca do Congresso Americano:

Pode-se dizer que com os avanços das tecnologias, o MARC vem novamente se adaptando as novas tendências mostrando que a união entre a estrutura de representações bibliográficas e catalográficas MARC e a linguagem XML será extremamente valiosa. (FLAMINO, 2006, p. 120).

Um exemplo é a utilização do formato MARC para classificação, que é discutido por Ford (2013, p. 168) em sua pesquisa, e cuja ênfase recai nas revisões sobre os cabeçalhos de autoridade e nas estruturas hierárquicas providas pelos

vocabulários descritivos *Simple Knowledge Organization System* (SKOS) e RDF. Sua análise sugere a construção de ontologia para a Classificação Decimal de *Dewey* (CDD) a partir da *Online Computer Library Center*(OCLC),e sua integração por meio de *linked data* aos cabeçalhos de autoridade.

A partir dessas considerações, e sem prejuízo de inúmeros estudos de caso e revisões de literatura que contemplam o universo da descrição bibliográfica no ambiente web, percebe-se a reformulação do papel descritivo dos metadados digitais. Sua aplicabilidade não se limita a reproduzir os indicadores, campos e subcampos do formato MARC; mas em preparar a descrição bibliográfica para o cenário multiplataforma, cujos paradigmas relacionais emergentes como os adotados pelo FRBR, figuram como estratégias para viabilizar a integridade e o acesso aos recursos bibliográficos.

É neste aspecto que o *joint steering committee* (JSC), ao tratar do esforço colaborativo para o desenvolvimento do RDA, ressalta que seu desenvolvimento reconhece que as bibliotecas agora operam em um ambiente digital baseado na web, o que envolve relacionamentos com criadores e usuários de metadados fora do espaço da biblioteca. O RDA apoiará a integração dos registros catalogados nas bibliotecas com aqueles produzidos por outras comunidades de metadados, fazendo com que os registros das bibliotecas sejam utilizados de novas maneiras (JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA, 2012).

Ao ressaltar o Seminário sobre Registros Bibliográficos realizado em Estocolmo no ano de 1990, os FRBR destacam dois pontos de particular relevância: a redução dos custos da catalogação e a promoção de estratégias capazes de ir ao encontro das necessidades dos usuários que lidam com variados suportes de informação. Este cenário repercute diretamente no uso dos metadados e nas estratégias voltadas a conversão para ambientes multiplataforma, aspectos que contemplam as demandas identificadas pela IFLA e relatadas na literatura.

Abordar a conversão de metadados como estratégia para o domínio bibliográfico possibilita a identificação de fatores que interferem na consistência, precisão e integridade dos dados. Algumas técnicas têm auxiliado sobremaneira esta tarefa, como a *crosswalking* tratada no capítulo 2.

Esta pesquisa ressalta tendências emergentes para a Biblioteconomia e Ciência da Informação no que se refere a adequabilidade de estruturas de vocabulários

descritivos e esquemas de metadados no âmbito da web, particularmente no tratamento de formas autorizadas de nomes pessoais e assuntos

1.3 Objetivo Geral

Pesquisar estratégias de conversão dos dados de autoridade de nome e assunto do formato MARC21 a fim de representá-los nos esquemas de metadados MARCXML, MADS e MODS com ênfase na redução de perdas de informação e enriquecimento lógico-semântico baseado na estrutura conceitual dos FRBR.

1.4 Objetivos específicos

- 1 Mapear os campos de autoridade de nome e assunto do formato MARC bem como as entidades definidas no FRBR;
- 2 Correlacionar os campos com propriedades e atributos dos metadados digitais MARCXML, MADS e MODS da *Library of Congress*;
- 3 Identificar a representação dos recursos catalogados em MARC21 nos metadados digitais acima a partir do modelo entidade-relacionamento dos FRBR.

1.5 Estrutura da dissertação

Esta dissertação está estruturada em 8 capítulos. No primeiro, apresentamos a introdução que traz aspectos gerais da catalogação e o papel de algumas instituições na sua evolução ao longo do tempo. Em seguida, na forma de subtópicos, são tratados: a problemática, justificativa e os objetivos gerais e específicos da pesquisa.

A etapa metodológica consta no capítulo 2, com as diretrizes, métodos e técnicas da pesquisa, bem como o percurso das etapas envolvidas na análise empírica.

O referencial teórico da pesquisa está no capítulo 3 e inicia com a conceituação, origem e aplicabilidade dos metadados, para, em seguida, introduzir a discussão do formato MARC. Nos tópicos seguintes foram apresentadas as normas e padrões de codificação e controle empregados na validação do formato MARC e dos esquemas de metadados descritivos em estudo.

O capítulo 4 e subtópicos destinam-se aos esquemas de metadados METS, MARCXML, MADS e MODS. Abordam seus conceitos e estruturas em estreita relação com o capítulo 5 seguinte, que contemplam a modelagem conceitual do domínio bibliográfico, momento em que são apresentados os conceitos atinentes aos requisitos

funcionais para descrição bibliográfica, notadamente aqueles dos grupos 2 e 3 dos FRBR, entidades de autoridade e assuntos respectivamente.

O capítulo 6 é dedicado à parte empírica da pesquisa, e foi dividido em duas seções, que contemplam a tarefa de conversão dos pontos de acesso da descrição. A princípio, o mapeamento das *tags* de autoridade de nomes e assuntos a partir dos campos e subcampos correspondentes do formato MARC e sua posterior codificação nos esquemas de metadados digitais.

Análises detidas sobre as propriedades dos esquemas MADS e MODS bem como as funções por elas exercidas foram incorporadas ao paradigma relacional, como opção metodológica recorrente em todas as etapas do mapeamento.

Nos capítulos da etapa empírica destinados às codificações nos esquemas XML aqui adotados, quais sejam capítulos 6.3, 6.4 e 6.5, os atributos e entidades do modelo conceitual dos FRBR, tratados no capítulo 5, foram retomados como instâncias e propriedades dos metadados MARCXML, MADS E MODS.

Por fim, na discussão dos resultados no capítulo 7, consta a análise dos achados na etapa empírica. Já no capítulo 8 estão as conclusões sugestões de pesquisas futuras que versem sobre conversão de metadados no contexto da catalogação, com discussões baseadas no conteúdo das revisões de literatura e nos resultados alcançados.

2 DIRETRIZES METODOLÓGICAS: MÉTODOS, TÉCNICAS E UNIVERSO DA PESQUISA

Os estudos exploratórios e descritivos correspondem às etapas iniciais desta pesquisa, propiciados por levantamentos bibliográficos e documentais, posto que esse tipo de pesquisa tem o condão de nos permitir compreender melhor o objeto de estudo e expor de maneira coerente as características do fenômeno estudado.

Segundo Marconi e Lakatos (2016, p.80), o estruturalismo é uma abordagem científica que pretende descobrir a estrutura do fenômeno, penetrar em sua essência para determinar as suas ligações determinantes. O método nasceu nos primeiros anos do século XX, representando uma crítica aos postulados positivistas e evolucionistas. (TRIVINOS, 2013, p. 120). Esta perspectiva dinâmica da estrutura integrante de vocabulários e esquemas descritivos, rompe com a linearidade e univocidade outrora vigente, abrindo novos horizontes de representação da informação registrada.

Neste contexto, e em consideração a organização hierárquica dos atributos e das propriedades dos esquemas descritivos tratados na pesquisa que tem no paradigma relacional uma característica recorrente, o método estruturalista torna-se o mais adequado para tratar as relações entre atributos e propriedades descritivas no contexto das implicações das partes com o todo e vice-versa.

Paralelo ao método estruturalista adotou-se a observação sistemática como técnica de pesquisa, em consideração à relevância da modelagem conceitual e do paradigma entidade-relacionamento e sua integração ao corpus da pesquisa. Sobre esta técnica, Marconi e Lakatos (2016, p. 78) apontam três designações que a constituem: a estruturada, a planejada e a controlada. Para os objetivos desta pesquisa, adotou-se a primeira opção, que viabiliza a análise conjunta dos esquemas, sintaxes, propriedades e as suas relações no âmbito do domínio bibliográfico. Ainda segundo Richardson (1985, p. 150), a adoção da referida técnica somente será possível quando se tem algum conhecimento do problema, uma vez que se viabiliza o estabelecimento de categorias em função das quais se deseja analisar a situação. As categorias concernentes a esta pesquisa serão detalhadas a seguir.

Para atendimento dos objetivos da pesquisa serão exploradas as ferramentas e os procedimentos aplicados em estudos de caso para conversão de metadados. Sobre

este ponto, e como diretriz para mapeamento e extração dos atributos e valores dos esquemas MARCXML, MADS e MODS em análise, Almeida e Cendon (2013, p. 16) apontam dois aspectos recorrentes para esta tarefa relativos às ferramentas de conversão. De um lado existem as ferramentas de conversão direta, destinadas à construção das marcações de um documento XML a partir de uma equivalência direta dos campos e subcampos do formato MARC. De outro, o procedimento que considera a estrutura do documento XML como reflexo dos significados dos campos do padrão MARC. Uma vez que o modelo entidade-relacionamento adotado pelos FRBR será o paradigma norteador da análise dos esquemas MARCXML, MADS e MODS, ambos os aspectos serão adotados.

Foram delimitadas duas categorias de registros em MARC, divididas em dados de autoridade de nomes e assuntos, contemplando os grupos 2 e 3 do FRBR respectivamente.

Posteriormente, com o uso da ferramenta de conversão *MARC Editor*, proceder-se-á a conversão dos arquivos *.mrc* para a extensão *.xml*, salvando-os na estrutura dos esquemas de metadados em estudo. A técnica *crosswalking*, será o modelo de mapeamento adotado para fins de identificação e anotação dos limites e possibilidades dos esquemas-alvo da conversão com vistas ao atendimento dos objetivos propostos.

O modelo *crosswalking* é destinado a mapear elementos sintáticos e semânticos de um esquema de metadados para outro. (NISO, 2004, p. 11). Para Assumpção (2013, p.73), esta técnica ainda possibilita o estabelecimento dos vínculos entre metadados semanticamente equivalentes em padrões distintos.

O cotejamento entre campos-alvo e campos-fonte dos esquemas de metadados, será representado por diagramas entidade-relacionamento em correlação às entidades dos grupos 2 e 3 do FRBR. O conceito entidade-relacionamento, formulado por Peter Chen em 1976, é baseado em níveis que objetivam proporcionar uma visão lógica da estrutura de dados em um modelo. Esses níveis compreendem:

- Informações relativas a entidades e relacionamentos que existem em nossas mentes;

- Estrutura de informação - organização da informação nas quais entidades e relacionamentos são representados por dados;
- Estruturas de dados independentes de acesso – as estruturas de dados que não são envolvidas com esquemas de busca, indexação e etc;
- Estruturas de dados dependentes. (CHEN, 1976, p. 10)

O modelo em análise compreende os dados de autoridade de nomes pessoais e assuntos tópicos, que no formato MARC correspondem aos campos 100 e 650 respectivamente. Toda tarefa de *crosswalking*, análise dos dados obtidos, conversão, e elaboração de modelos, na perspectiva da observação sistemática adotada, tem como foco os dois aspectos constantes do objetivo geral da pesquisa: identificação das variáveis que resultam em perdas de informação e as possibilidades para incremento lógico-semântico do modelo conceitual.

O mapeamento das entradas que envolvem assuntos tópicos, além do campo 650, inclui também outras variantes, quais sejam: campo 600 – nome pessoal como assunto, 610 – entidade, 611 - eventos e 651 nome geográfico. O levantamento dessas formas variantes de entradas de assunto é imprescindível, na medida em que possibilita explorar as estratégias relacionais de atributos presentes nos esquemas de metadados digitais MARCXML, MADS e MODS em estudo.

3 SOBRE METADADOS DESCRITIVOS

Os metadados destinados à representação de objetos de informação podem tanto ser aplicados à catalogação propriamente dita, como podem exercer a função de protocolos de comunicação. Exemplos de protocolos e perfis de metadados aplicados à definição arquitetural de sistemas web são: *Simple Object Access Protocol* (SOAP), *Representational State Transfer* (REST), arquitetura orientada a serviços (SOA) e o *Open Archives Initiative* (OAI-PMH).

Embora padrões de metadados para a *web* sejam codificados para atendimento a finalidades específicas, tais como serviços em rede como o caso do SOAP, eles possuem sua sintaxe baseada na linguagem XML. Esta composta por *namespaces* que contemplam propriedades, atributos e valores, em atendimento aos requisitos de um documento bem formado, apto a ser processado por humanos e máquinas que se utilizam de analisadores sintáticos ou *parsers*.

A elevada demanda por padrões e formatos, por consequência, tem conduzido a um grande número de esquemas. Segundo Zeng (1999, p. 1194) duas categorias contemplam o universo de metadados disponíveis: os gerais e especializados. A classificação dessas categorias é definida por algumas variáveis que envolvem o conteúdo do recurso a ser descrito, o usuário do recurso e o recurso propriamente dito.

O papel descritivo dos metadados pode envolver atributos textuais verbais e não verbais a exemplo do formato *Visual Resources Association Core Categories for Visual Resources* ou *VRA Core*, destinado a representar objetos visuais de acervos culturais.

Abordar os metadados para catalogação, segundo os novos paradigmas ensejados pelas tecnologias digitais de informação e comunicação, demanda uma análise de seu conceito, origem e aplicabilidade antes de adentrarmos ao domínio bibliográfico. Visto que seu uso e relevância é objeto de estudos por diversas áreas e especialidades, bem como os formatos disponíveis tratados neste capítulo.

3.1 Origem

A palavra metadado é mais uma que aparece no meio de tantas outras introduzidas e divulgadas pelas Tecnologias Digitais de Informação e de Comunicação

(TDIC). Porém, sua gênese está nas reflexões de Philip Bagley, no ano de 1968¹, cuja origem é constituída por dois termos: *meta*, prefixo grego (μετά) que exprime a idéia além de, ou ainda, traz a semântica de abstração como podemos ver em metafísica e meta linguagem. Por sua vez a palavra dado está relacionada à representação objetiva ou elementar de uma realidade.

A definição de metadados como "dados sobre dados" representa um consenso geral na literatura das áreas que lidam com organização e tratamento da informação registrada. (ALVES, 2010; BAKER, 2012; BEZERRA, 2013; BUCKLAND, 2006; FUSCO, 2010). Neste aspecto, El-Sherbini e Klin (2004, p. 239) ressaltam que o entendimento geral orientou a ampliação de seu escopo conceitual para lidar com a descrição bibliográfica padronizada, o que inclui inúmeras variantes de suportes e modalidades de acesso aos recursos informacionais.

Essa desvinculação do suporte e primazia no conteúdo, encontrou no ambiente distribuído um cenário emergente de ressignificação de conceitos e redefinição de prioridades. Embora a finalidade descritiva dos metadados seja uma constante em paralelo à relevância da adoção de regras e padrões para descrição e intercâmbio.

Dos padrões de metadados mais relevantes e estudados na atualidade estão aqueles que se inserem no ambiente *web*. Ao assumir papel relevante para organização, tratamento e recuperação da informação, Baker (2012, p 118) assinala como padrão pioneiro o DC, discutido a seguir, por representar um "vocabulário chave" no contexto da web semântica e dos dados lincados.

Ao incorporar o papel dos metadados na padronização proporcionada pelas regras da catalogação, importante considerar que:

Novos e antigos metadados são baseados em práticas comuns. Padrões de catalogação foram criados e desenvolvidos como forma de organizar a informação para fins de facilitar sua localização e acesso. Os padrões são os fundamentos sobre os quais todas as regras de metadados e catalogação são

¹Philip Bagley « As important as being able to combine data elements to make composite data elements is the ability to associate explicitly with a data element a second data element which represents data "about" the first data element. This second data & ement we might term a "metadata element". Examples of such metadata (elements are: an identifier, a domain "prescriptor" which specifies from what domain the values of the first element must be taken, an access code which limits the conditions under which the first data element c-in be accessed. »,

BAGLEY, Philip. Extension of programming language concepts, Philadelphia, University City Science Center, 1969, p. 26. Disponível em: <<http://stevenbagley.net/blog/philip-r-bagley.html>> Acesso em 20 de nov. 2014.

desenvolvidos. Sem esses padrões para catalogação, um simples item seria catalogado diversas vezes e cada registro catalogado conteria informações distintas [...] sem a existência de padrões para catalogação seria difícil imaginar como pesquisadores acessariam recursos e como as bibliotecas compartilhariam registros. (EL-SHERBINI; KLIN, 2004, p. 240)

Em complemento ao exposto acima, visualizamos os padrões de metadados como parte de um todo, em que os esquemas de descrição por eles proporcionados não atingiriam suas finalidades se estivessem desprovidos de protocolos de comunicação para fins de efetivar o intercâmbio, se não houvesse consenso da comunidade ao adotarem padrão específico e se os desenvolvedores de soluções tecnológicas não aderissem às recomendações formais destinadas à uniformização de dados.

Para a web semântica, as recomendações do consórcio *w3c* são mandatórias, visto que sem elas, os requisitos supracitados não poderiam cumprir em larga escala o compartilhamento, identificação, acesso e uso de recursos informacionais. Conforme se verifica no grupo de trabalho do RDF, OWL, XHTML somente para citar alguns.

Estas características dos metadados são vistas também no domínio bibliográfico, em particular, na revisão das regras de catalogação e na adaptação das sintaxes das estruturas do formato MARC às linguagens XML. Este aspecto resulta em novos desafios para a codificação dos registros bibliográficos, como aqueles relativos à conversão de formatos.

3.1.1 Aplicabilidade

Nos anos de 1960 Myers (1960 *apud*MCCALLUM, 2004, p. 85), definiu metadados como "conjunto de dados" destinado a finalidades comerciais e com amplo potencial descritivo. Posteriormente, os padrões de metadados passaram a ser um importante instrumento para armazenar, organizar, descrever e prover acesso às fontes de informação para diversas especialidades.

Nas décadas de 80 e 90 respectivamente, os metadados ganharam nova dimensão. Segundo Alves (2010, p. 44) nos anos 80 foram incorporados aos Sistemas de Gerenciamento de Bases de Dados que os empregavam na descrição das informações contidas nas bases. Nos anos 90, por meio de seu uso por diversas comunidades para padronização e interoperabilidade, com destaque para o setor geoespacial, considerado a comunidade pioneira em seu uso.

Atualmente, verifica-se aplicabilidade dos padrões de metadados na área da educação, a exemplo do *Learning Object Metadata* (LOM), no campo da saúde através do *Unified Medical Language System* (UMLS), na área da matemática por meio da especificação do *World Wide Web Consortium* (W3c) denominada MathXML, dentre outros.

Ainda nos anos 90, Zeng (1999, p. 1198) assinala o esforço interdisciplinar e internacional para definição de um conjunto de elementos voltados à catalogação de objetos digitais, que culminou no advento do DC. O autor o define em três grandes grupos: o que contempla elementos relativos ao conteúdo dos recursos, aos dados de autoridade intelectual e as instancias do recurso.

A origem do *Dublin Core* (DC) decorreu de um workshop realizado em Dublin Ohio no mês de março de 1995, motivado pela questão: "Pode ser definido um registro de metadados simples que descreva uma vasta gama de objetos eletrônicos?" Cole(2002, p. 84). Um repertório terminológico objetivo e consensual para descrever objetos digitais foi criado, e recebeu a designação de *Dublin Core Metadata Element Set*, composto por 15 elementos. As limitações descritivas do DC são abordadas por Baker (2012, p.118), que referencia diversos encontros e debates em torno da contribuição do DC para os registros catalogados no formato MARC, e a necessidade da integração com pacotes e estruturas mais complexas, a exemplo do RDF.

Os 15 elementos que formam o núcleo do DC são comumente incorporados a outros esquemas de metadados, tais como XML e RDF. O objetivo da integração é proporcionar meios para o alcance da interoperabilidade bem como a formalização de uma semântica adequada para descrever diferentes objetos informacionais. (ALVES, 2010; CATARINO; SOUZA, 2012; GONZALES, 2014; SOUZA;BEZERRA; COSTA, 2016).

3.1.1.1 Formato MARC

Padronizar registros bibliográficos com ênfase na redução dos custos da catalogação e na promoção da interoperabilidade corresponde a um esforço de longo prazo efetivado por comitês e instituições ao redor do mundo. Segundo Assumpção (2013, p. 21), com o advento dos formatos legíveis por máquina na década de 60 pela Biblioteca do Congresso Americano, destinados à transmissão de registros em fitas magnéticas, surgiu o caráter sequencial da codificação original do formato utilizado ainda hoje.

O formato MARC, posteriormente chamado de USMARC, passou por muitas modificações desde o projeto inicial. Outros países, segundo Moreno e Brascher (2007, p. 15), adaptaram o formato e criaram modelos próprios a partir do MARC, como, por exemplo, o CAN/MARC no Canadá; o MONOCLE na França; FINMARC, na Finlândia; IBERMARC na Espanha e a Catalogação Legível por Computador (CALCO) no Brasil. A adoção em diversas realidades nacionais de formatos distintos conduziu a inconsistências com potencial para afetar a interoperabilidade. Para superar esta lacuna, foi criado o UNIMARC:

O formato UNIMARC surgiu como uma tentativa de resolver o problema de incompatibilidade. Foi desenhado de uma tal maneira que qualquer registro MARC nacional pudesse ser razoavelmente representado no UNIMARC sem distorções severas. A intenção foi torná-lo intermediário para conversão entre dois formatos MARC. (SHMIDT; PATEL, 1999, p. 287)

Segundo Aalberg e Zumer (2012, p. 852), é a partir dos formatos UNIMARC e MARC21 destinados a cumprir o papel de intercambio de registros bibliográficos, que outros formatos foram derivados, a exemplo daqueles citados no parágrafo anterior. Ressaltam os autores a primazia do MARC 21 como formato adotado internacionalmente, não apenas para intercambio dos registros, mas também para estruturar, armazenar, gerenciar e usar dados.

Quanto a sua estrutura, o formato MARC é composto pelo campo líder, que compreende informações do registro como um todo e dos campos e subcampos que armazenam conteúdo atual do registro. O campo líder pode identificar os tipos de registros (livros, sons, gravações, materiais cartográficos dentre outros), nível bibliográfico (monografias e séries) e outras informações genéricas do registro.

A norma responsável pela definição dos campos líder, diretório e das variáveis do MARC é a ISO 2709, que prescreve a sintaxe do formato destinada à padronização da estrutura do registro bibliográfico MARC, para que este se torne apto a interoperar em ambiente de rede.

Os campos de autoridade e assunto tratados nesta pesquisa contemplam os indicadores e sinalizadores. Segundo Almeida e Cendon (2003, p. 4), os indicadores, que são representados pelos caracteres constantes no quadro 2, bem como os sinalizadores, delimitam os campos e subcampos do formato e se destinam à leitura por máquina. Para as autoridades de nome e assunto, o quadro traz a função que os indicadores exercem na forma de entrada do ponto de acesso:

Quadro 2 – Forma da entrada autorizada de nome e assunto segundo indicadores.

Campo	Autoridade	Valor dos Indicadores	Forma de entrada
100	Nome pessoal	0	Prenome
		1	Sobrenome
		2	Nome de família
110	Entidade	0	Nome invertido
		1	Nome da jurisdição
		2	Ordem direta
650	Assunto tópico	#	<i>Undefined</i>
		0	Nível não especificado
		1	Primário
		2	Secundário
610	Assunto –Entidade	0	Nome invertido
		1	Jurisdição
		2	Ordem direta

Fonte: Elaborado pelo autor.

O indicador representado pelo caractere 0 no campo 100 significa que a entrada da autoridade de nome dar-se-á pela sua ordem direta, ou seja, prenome seguido do sobrenome. Já o valor do indicador 1 do mesmo campo trata da forma invertida, cuja entrada será pelo sobrenome, e o valor do indicador 2 indica que a entrada autorizada será pelo nome de família: dinastia, clã ou família.

Os demais campos possuem funções distintas definidas pelos indicadores, conforme a posição que devem assumir no catálogo ou nas bases de autoridades. Já o sinal # indica que a informação não está definida, ou *undefined*, conforme convenção notacional da LC. Na figura 1 temos a representação de uma autoridade de nome pessoal com seus respectivos indicadores e sinalizadores conforme descrito acima.

Figura 1 – Representação de indicadores e sinalizadores em campo de autoridade de nome pessoal.

100	1	\$a	Cassirer, Ernst	\$d	1874-1945
-----	---	-----	-----------------	-----	-----------

Fonte: Elaborado pelo autor.

O subcampo \$a e subcampo \$d na figura 1, são notações definidas pela ISO2709 relativa a dados de formato legível por máquina. O primeiro subcampo é atribuído ao nome pessoal e o segundo à data. Os sinais e símbolos do formato MARC

carecem da semântica necessária para sua legibilidade por humanos, aspecto adequadamente tratado com a linguagem XML.

3.1.1.2 Linguagem XML

O consórcio *w3c*, organização vanguardista na criação e recomendação dos vocabulários integrantes das tecnologias baseadas na web semântica, criou em 1997, a linguagem *Extensible Markup Language* (XML). O consórcio a define como uma sintaxe simples e flexível derivada de seu predecessor *Standard Generalized Markup Language* (SGML), e surgiu para fazer frente às publicações eletrônicas em larga escala (W3C, 2016). Quanto ao contexto de uso, Lam (2002, p. 90) a define como um padrão simplificado da SGML, que permite preparar o texto para transportar não apenas informação de layout, mas também estrutura semântica.

Em ambientes abertos e distribuídos como a Internet, onde um catálogo pode ser acessado a partir de diversos locais, as linguagens de marcação são a principal forma de representar e transportar dados. Nesse contexto, destaca-se a linguagem XML (*Extended Markup Language*), a qual é uma linguagem de marcação publicada em 1997 pelo W3Consortium, que tem aplicação na descrição de conteúdo e transporte de dados. (ALMEIDA; CENDON, 2003, p. 2)

Ainda no que concerne ao uso de documento XML e sua relevância para as bibliotecas, Vidakovic e Rackovic (2006, p. 514) o considera adequado para sistemas de informação em bibliotecas, uma vez que seus documentos são estruturados e o uso da tecnologia XML é apta para modelá-los.

O papel atribuído ao XML para modelagem de dados também é tratado por Flamingo (2006, p. 107), ao afirmar que é por meio desta característica que se torna possível definir quais elementos o documento terá e os vínculos entre eles.

A sintaxe do XML é representada a seguir. As principais seções do documento estão sinalizadas, sendo que a declaração no prólogo inclui a codificação latina de caracteres ISO-8859-1. É comum encontrarmos nesta declaração o padrão *8-bit Unicode Transformation Format* ou UTF-8 que traz como características ser um formato *unicode* e de tamanho variável, conforme a figura 2:

Figura 2: Sintaxe de um documento XML.

```

<?xmlversion="1.0" encoding="ISO-8859-1"?> Declaração
<livro>
<titulo>Linguagem e Imagem</titulo>
<autor>José Sena Filho</autor>           Miscelânea
<assunto>Publicidade e Propaganda</assunto>
<editora>L & A editora</editora>
<ano>2004</ano>
</livro>

```

Fonte: Elaborado pelo autor.

A seção denominada declaração ou prólogo informa ao sistema que se trata de um documento XML com a codificação ISO8859-1. A miscelânea, por sua vez, expressa a flexibilidade tratada anteriormente na criação de marcações pelo próprio usuário, com o auxílio de *namespaces* que comportam propriedades e atributos destinados a representar um recurso de informação.

Diversas pesquisas são desenvolvidas com o objetivo de explorar o potencial da linguagem para representar e recuperar informações na web. Com este propósito, Taheri, Hariri e Fattahi (2014, p. 798) avaliaram a indexabilidade e recuperabilidade de documentos catalogados em MARCXML em motores de busca. Os resultados mostraram que mesmo na ausência da declaração no prólogo do documento XML e na exclusão da extensão do arquivo, os níveis de precisão e cobertura temática se mostraram idênticos aos documentos catalogados com essas estruturas sintáticas. Na análise dos pesquisadores, os analisadores sintáticos dos mais avançados sistemas de buscas estão aptos a ler e interpretar a hierarquia das *tags* e *namespaces*, mesmo na ausência de alguns metadados.

Outro atributo importante da linguagem XML é sua capacidade de representar e interoperar dados na rede, em particular benefício às plataformas de acesso aberto, tais como os repositórios institucionais, a exemplo da plataforma *dspace* que adota o protocolo OAI-PMH para *harvesting*.

Uma vez que os próprios criadores de esquemas de metadados em XML definem quais atributos e tipos de dados deverão constar da descrição, regras de validação são mandatórias. Os capítulos seguintes tratam das especificações para validação de esquemas aplicados ao domínio bibliográfico, com ênfase particular ao ambiente web.

3.2 Especificações de codificação e controle

O controle sintático das estruturas que integram os esquemas de metadados digitais baseados na XML é exercido pela Definição de Tipos de Documento (DTD). Em paralelo, protocolos e normas específicas como as analisadas neste capítulo, contribuem com a persistência e integridade dos dados para efetivação da interoperabilidade.

Os protocolos, segundo Flaminio (2006, p. 89), correspondem a um conjunto de regras que governam a comunicação entre diferentes plataformas computacionais, em alusão aos protocolos usados na sociedade para administrar as relações humanas.

O protocolo Z39.50 ao promover o acesso simultâneo a catálogos públicos de acesso online (OPAC), incorpora em termos gerais, a mesma finalidade do *Transmission Control Protocol/Internet Protocol* (TCP-IP), neste caso, para estabelecer a necessária comunicação entre computadores como processo viabilizador da internet.

Sobre este aspecto, Gartner (2014, p. 297) atribui aos dados legíveis por máquina a propriedade de promover o intercâmbio e a busca federada, o que implica abordar criticamente o papel do formato MARC em conjunto com as especificações correntes, destinadas a operacionalizar o intercâmbio e acesso aos recursos de informação na ambiência multiplataforma.

Tais especificações relevantes para a descrição bibliográfica em plataformas web serão tratadas a seguir.

3.2.1 Document Type Definition (DTD)

A DTD agrega convenções sintáticas cuja finalidade precípua é validar a estrutura de um documento XML. É um arquivo que pode tanto ser externo quanto interno a um documento XML Flaminio (2006, p. 108), com declaração no prólogo do arquivo conforme modelo apresentado na figura 3:

Figura 3: Sintaxe de um DTD interno.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!DOCTYPE livro [
<!ELEMENT livro (autor, titulo, editora)>
<!ELEMENT autor (#PCDATA)>
<!ELEMENT titulo (#PCDATA)>
```

```

<!ELEMENT editora (#PCDATA)>
] >
<livro>
<autor>M</autor>
<titulo>TutorialsPoint</titulo>
<editora>(011) 123-4567</editora>
</livro>

```

Fonte: Elaborado pelo autor.

A propriedade *standalone* possui o atributo *yes* sempre que a estrutura do DTD for interna, ou seja, estiver presente no corpo do documento XML. Quando o arquivo DTD for externo utiliza-se o atributo *no*, seguido da especificação da localização do arquivo *.dtd* para que o mesmo seja incorporado ao documento. O nome do arquivo é uma propriedade da *tag DOCTYPE*, conforme modelo da figura a seguir.

A cláusula *#PCDATA*, ou *Parsed Character Data*, é destinada a declarar sequências de elementos em um arquivo XML, as eventuais propriedades que lhes são intrínsecas e como o *parser* deve efetuar sua leitura. A figura 4 traz um exemplo de arquivo de validação incorporado ao documento XML.

Figura 4: Sintaxe de um DTD externo.

```

<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<!DOCTYPE address SYSTEM "address.dtd">
<livro>
<autor>Tanmay Patil</autor>
<titulo>TutorialsPoint</titulo>
<editora>(011) 123-4567</editora>
</livro>

```

Fonte: Elaborado pelo autor.

Uma vez que um documento XML tem como característica a sua flexibilidade notacional o que o torna apto a ser produzido e editado por seus autores, Almeida e Cendon (2003, p. 8) apontam regras sintáticas que integram um XML bem formado: a existência de um único elemento raiz, atributos delimitados por aspas e declaração explícita das entidades.

O papel da DTD no controle da estrutura de um documento XML destinado a representar registros bibliográficos definiria, por exemplo, a ordem dos campos, subcampos e indicadores de um registro MARC. Sobre as duas primeiras DTD's criadas pela LC, Assumpção e Santos (2013, p. 112) destacam que a primeira se destinava aos registros bibliográficos e informações para a comunidade; e a segunda aos registros de autoridade e de classificação.

As regras de validação do esquema são discutidas por Keith (2003, p. 122) que ressalta o tamanho considerável das duas DTD's resultantes e eventuais dificuldades a serem enfrentadas nas aplicações. Segundo Moreno e Brascher (2007, p. 18), as duas DTD's foram inicialmente desenvolvidas para a SGML e posteriormente convertidas para XML integradas ao MARCXML.

A SGML é tratada pelos autores como uma linguagem para descrever a estrutura dos documentos e promover o intercâmbio. Sua origem remonta a 1986 com a publicação da norma ISO 8879 *Information processing - Text and office systems* Assumpção (2013, p. 40), e deu origem a XML.

A criação do MARCXML *Schema* surgiu como alternativa viável para o MARCXML DTD, na medida em que passou a tratar adequadamente os metadados de um registro ISO2709, o que envolve seis elementos, segundo Keith (2003, p. 123): coleção, registro, líder, campo de controle, campo de dados e subcampos.

3.2.2 XML schema

O XML *Schema*, em comparação ao DTD que o precede, potencializa a capacidade descritiva de um documento XML na medida em que amplia a estrutura de validação e controle de sua sintaxe W3 (2004). A estrutura da figura 5 ilustra um XML *Schema* simplificado que se destina a controlar os elementos de dados presentes no documento XML da figura 2:

Figura 5: Sintaxe XML Schema

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
<xs:element name="livro">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="titulo" type="xs:string"/>
      <xs:element name="autor" type="xs:string"/>
      <xs:element name="asunto" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:schema>
```

```
        <xs:element name="editora" type="xs:string"/> <xs:element name="
ano" type="xs:date"/>
    </xs:sequence>
</xs:complexType>
</xs:element>
</xs:schema>
```

Fonte: Adaptado de W3C (2017)

O nodo raiz do esquema XML da figura 5, atribui à propriedade *element name* o atributo livro, seguido pela propriedade subordinada *complexType* destinado a intercalar os demais atributos presentes na estrutura do documento XML representado.

Os elementos *xs:string* e *xs:date* são propriedades atribuídas aos tipos de dados presentes no *namespace* correspondente. De acordo com o W3C (2017), outros elementos qualificadores dos dados em XML *schema* são: *decimal*, *integer*, *boolean* e *time*.

A hierarquia inerente a esquemas XML também está expressa na estrutura de metadados aplicados à validação. A exemplo da figura 5, a referida hierarquia reproduz a do respectivo documento.

4 ESQUEMAS DE METADADOS APLICADOS À DESCRIÇÃO DE RECURSOS BIBLIOGRÁFICOS

Os capítulos seguintes contemplam os esquemas de metadados codificados para representar pontos de acesso da descrição. Para os objetivos desta pesquisa, a apresentação de cada esquema enfatizará as propriedades e atributos de seus *namespaces* orientados à incorporação de um registro descrito no formato MARC.

4.1 Metadata Encodings and Transmission Standard (METS)

O padrão para codificar e transmitir metadados de registros bibliográficos é definido e mantido pelo *Network Development and MARC Standards Office* da Biblioteca do Congresso Americano e recebe o acrônimo de METS.

Segundo Cundiff (2004, p. 53), trata-se de um esquema XML desenhado com o propósito de criar instancias de um documento XML que expressem a estrutura hierárquica de objetos digitais o que envolve seus nomes e localizações. Ainda segundo o autor, sua estrutura contempla as três categorias de metadados: o descritivo, administrativo e estrutural, bem como expressa sua sintaxe por meio do XML *Schema* tratado no capítulo 3.2.2 .

Ainda sobre sua estrutura, McDonough (2006, p. 149) enumera os seguintes aspectos considerados integrantes do esquema METS:

- Um cabeçalho sobre o próprio documento METS;
- Uma seção para metadados descritivos;
- Uma seção para metadados administrativos;
- Um inventário de arquivos de conteúdo compreendendo o objeto digital;
- Uma seção de lincagem estrutural;
- Um mapa estrutural;
- Uma seção de metadados comportamental.

O uso do METS para codificação de objetos digitais complexos, busca contemplar diversas facetas assumidas por tais objetos. A categoria de metadados comportamentais, por exemplo, visa identificar recursos dinâmicos necessários para acessar e visualizar um documento.

Embora destinado a descrição de objetos digitais, a sintaxe do esquema METS incorpora a flexibilidade do documento XML, o que permite destacar seções do esquema que representam os pontos de acesso que se deseja abordar:

O primeiro elemento em um METS é o elemento raiz <mets>. Um elemento raiz é requisito no XML. O elemento raiz METS tem 5 atributos XML

opcionais: ID e OBJID (Identificador de Objeto), como identificadores; TYPE especifica que tipo de objeto está sendo documentado; LABEL identifica o título do objeto; e PROFILE que especifica qual, se houver, perfil registrado no METS está sendo usado. O elemento raiz também contém declarações em namespace que declaram todos os esquemas de metadados usados no documento METS em conjunto com prefixos de namespaces associados. (WAGNER, 2011, p. 2)

Alguns autores como Guenther e McCallum (2003, p. 12) enfatizam a relevância do METS para viabilizar o uso e a preservação de objetos digitais, bem como sua flexibilidade no que concerne à representação de recursos de informação para bibliotecas digitais, com ênfase na capacidade descritiva superior a linguagens de marcação como a HTML. A aplicação do METS para compartilhamento de registros de informação é tratada por Ralson (2013, p. 2), que além de ressaltar estes pontos, alerta para seu uso no contexto da acessibilidade. Afirma o autor que:

Esquemas de metadados apropriados para uma dada comunidade de usuários provêm frameworks e métodos para manipulação de sintaxes e semânticas que ajudam a padronizar as formas como a informação associada a objetos – ambos simples e complexos – são comunicadas. (RALSON, 2013, p. 3)

Uma das facetas do METS é a flexibilidade para incorporar outros esquemas de metadados. Como exemplo, destaca-se o *namespace* <mdWrap> que a LC denomina por Metadados Descritivos Internos. Trata-se de um container cuja finalidade consiste em incorporar sintaxes de outros metadados como o DC conforme a figura 6:

Figura 6: Sintaxe do esquema METS.

```
<dmdSec ID="dmd002">
  <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Dublin Core
  Metadata">
    <xmlData>
      <dc:title>Alice's Adventures in Wonderland</dc:title>
      <dc:creator>Lewis Carroll</dc:creator>
      <dc:date>between 1872 and 1890</dc:date>

      <dc:publisher>McCloughlin Brothers</dc:publisher>
      <dc:type>text</dc:type>
    </xmlData>
  </mdWrap>
</dmdSec>
```

Fonte: LIBRARY OF CONGRESS (2017)

O atributo DC da propriedade *mdtype* sinaliza ao analisador que os metadados intercalados pela *tag xmlData* são provenientes de um vocabulário *Dublin*

Core. A propriedade *label*, por sua vez, destina-se a fornecer um título não abreviado com a finalidade de permitir a visualização do metadado incorporado ao esquema.

As propriedades de um esquema METS, em virtude de seu caráter genérico, têm o condão de preparar o objeto a ser catalogado em conformidade com outros padrões e normas para descrição. Nesta pesquisa, os dados em análise correspondem às entradas de nomes e assuntos, o que demanda a adoção de esquemas de metadados descritivos alinhados às regras para descrição bibliográfica. Neste aspecto, o esquema MARCXML não apenas incorpora o subconjunto dos elementos MARC numa arquitetura flexível e interoperável, como permite abordar de forma consistente os atributos da descrição e inconsistências nos vínculos e remissivas no interior do registro bibliográfico conforme demonstrado no capítulo 6.

4.2 MARCXML

Desenvolvido pelo *Library of Congress' Network Development and MARC Standards Office* em 2002, o MARCXML incorpora a flexibilidade e extensibilidade da linguagem XML, permitindo ao usuário trabalhar o formato MARC numa sintaxe capaz de proporcionar o acesso e a recuperação de registros bibliográficos no contexto multiplataforma.

A estrutura deste esquema é derivada da sintaxe XML *Schema* tratada no capítulo 3.2.2, que segundo Flaminio (2006, p. 107), expressa vocabulários compartilhados e permite que máquinas executem regras elaboradas por pessoas. Provê meios para definir a estrutura, conteúdo e semântica de documentos XML.

De acordo com Assumpção (2013, p. 52), o surgimento da XML *Schema* levou a LC a criar uma nova forma destinada a codificar registros MARC 21 com a XML, surgiu assim o MARC 21 XML *Schema*, também conhecido como MARCXML.

Como se trata de um esquema com marcação derivada do subconjunto de campos do formato MARC e orientada ao ambiente web, Moreno e Brascher (2007, p. 18) citam a relação de tecnologias adicionais que integram seu escopo, tais como *stylesheets* (folhas de estilo), *softwares* e ferramentas de conversão. Estes componentes são vitais para o processo de conversão, com particular ressalva ao XML *schema* e as folhas de estilo destinadas a controlar a exibição dos dados na interface com o usuário.

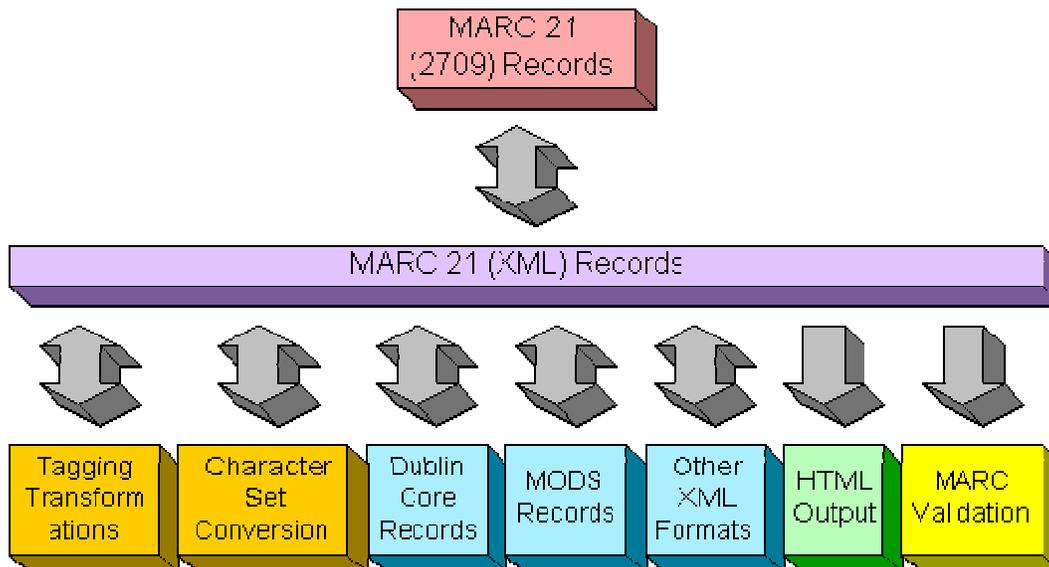
A implementação do MARCXML em bibliotecas digitais e coleções eletrônicas de dados tem sido tratada como fator estratégico para promover a interoperabilidade e acesso aos registros bibliográficos. A pesquisa desenvolvida por

Ochando (2014) ilustra como registros codificados em MARCXML podem ser adotados como substitutos de serviços de alerta e difusão de notícias, como o *Rich Simple Syndication* (RSS) para transportes de dados; e como grandes volumes de registros bibliográficos podem ser adequadamente representados.

No entanto, a LC situa o MARCXML em uma camada intermediária entre o registro MARC 21 e demais esquemas de metadados descritivos, cuja conversão para interface com o usuário é operacionalizada com a tecnologia denominada linguagem extensível para folhas de estilo de transformações ou *eXtensible Stylesheet Language for Transformation* (XSLT). (RILEY, 2010, p. 9).

O diagrama da LC que define a posição do MARCXML na escala de conversão de metadados é representado pela figura 7:

Figura 7: Posição do MARCXML na escala de conversão de metadados



Fonte: LIBRARY OF CONGRESS(2016)

Quanto ao uso do MARCXML, a Biblioteca do Congresso Americano estabelece os seguintes aspectos que trazem ao esquema a flexibilidade na incorporação a outros formatos e a interoperabilidade na comunicação no ambiente web:

- Representa uma extensão ao esquema METS tratado no capítulo 4.1;
- Representa um registro MARC completo em XML;
- Representa metadados para *harvesting*;
- Apto a descrição de recurso original em XML.

O subconjunto de campos e subcampos MARC encontram neste esquema, as propriedades e os atributos necessários capazes de expressar os pontos de acesso em

alinhamento notacional compatível com as áreas de descrição. Na figura 8, o MARCXML traz as áreas de título e assunto tópico que ilustram a flexibilidade do esquema em reproduzir na linguagem natural as propriedades que identificam o subconjunto de elementos MARC:

Figura 8: Áreas de título e assunto tópico em MARCXML.

```
<marc:datafield tag="245" ind1="0" ind2="4">
<marc:subfield code="a">The Great Ray Charles</marc:subfield>
<marc:subfield code="h">[sound recording].</marc:subfield>
</marc:datafield>

<marc:datafield tag="650" ind1=" " ind2="0">
<marc:subfield code="a">Jazz</marc:subfield>
<marc:subfield code="y">1951-1960.</marc:subfield>
</marc:datafield>
<marc:datafield tag="650" ind1=" " ind2="0">
<marc:subfield code="a">Piano with jazz semble.</marc:subfield>
</marc:datafield>
```

Fonte: LIBRARY OF CONGRESS (2017).

O *namespace* raiz do MARCXML da figura 8, destinado a representar os campos de dados, recebe a designação de *datafield* que associa à propriedade *tag* o valor corresponde ao campo MARC a ser representado. O *datafield* encapsula o elemento filho *subfield* que, por sua vez, trata dos subcampos do formato MARC, aqui representado pela propriedade *code*. Os dois indicadores do MARC estão presentes no *datafield* e representados pelos acrônimos *ind1* e *ind2* respectivamente.

O caráter extensível da linguagem XML, conforme assinalado anteriormente, permite que outros esquemas de codificação e representação sejam incorporados ao MARCXML, para fins de lidar com dimensões específicas. Na seção seguinte trataremos de uma dessas dimensões, o esquema MADS, que agrega semântica necessária para a representação de dados de autoridade.

4.3 Metadata Authority Description Schema (MADS)

No universo dos metadados para descrição bibliográfica, o MADS é um esquema XML destinado a codificar e representar um conjunto de elementos de autoridade que pode ser usado para prover metadados sobre agentes (pessoas, organizações) e eventos (GUENTHER, 2007, p 14). A mesma autora argumenta que o

MADS não tem a pretensão de substituir o formato MARC, mas tornar recursos de informação bibliográficos acessíveis na ambiência multiplataforma, bem como permite o estabelecimento de estratégias objetivas para tratamento de inconsistências na representação de entradas de autoridade de nomes e assuntos.

Os atributos do esquema MADS destacam do registro bibliográfico os pontos de acesso autorizados, ou, ainda, as formas autorizadas das entradas extraídas de vocabulários controlados como o da Biblioteca Nacional Brasileira e o *Library of Congress Subject Headings* (LCSH).

No que concerne a integridade do MADS para incorporar os dados de autoridade, Assumpção e Santos (2013, p. 121) argumentam que nem todos os campos, subcampos e indicadores do MARC 21 de autoridade possuem equivalentes no MADS, fato que pode ocasionar a perda ou a junção de dados após a conversão de um padrão para outro. Como consequência direta é verificada a necessidade de incorporar propriedades ou excluí-las conforme necessidade descritiva que não importe perda de informação.

Dentre os elementos considerados como principais que podem existir em um registro MADS estão: *authority*, *related* e *variant*. (ASSUMPCÃO; SANTOS, 2013, p. 117). Os autores argumentam que em um registro, é obrigatória a presença de, no mínimo, um elemento *authority*, já os elementos *related* e *variant* são opcionais.

O quadro 3 lista os subelementos dos respectivos *namespaces* citados no parágrafo anterior bem como as relações entre seus atributos no contexto da autoridade representada.

Quadro 3 - Atributos e subelementos dos principais namespaces do esquema MADS.

Namespace	Atributos	SubElementos
<authority>	ID, xlink, lang, xml:lang, script, transliteration geographicSubdivision (nenhuma, direta, indireta, não aplicável)	<name> <namePart> <description>
<related>	ID, xlink, lang, xml:lang, script, transliteration type (earlier, later, parentOrg, broader, narrower, equivalent, other)	<name> <namePart> <description>
	ID, xlink, lang, xml:lang, script, transliteration type (acronym,	<name> <namePart> <description>

<variant>	abbreviation, translation, expansion, other) (acrônimo, abreviatura, tradução, expansão, outra) otherType hidden (yes)	
-----------	---	--

Fonte: Elaborado pelo autor.

Os subelementos de *authority*, além dos mencionados no quadro 3, incluem diversos *namespaces* que objetivam contemplar todas as formas variantes de manifestação que uma autoridade poderá assumir. Dentre eles citamos: *titleInfo*, *topic*, *temporal*, *genre*, *geographic*, *hierarchicalGeographic* e outros. O uso dessas formas variantes ou facetas não é prescritivo, sua abordagem visa contemplar todas as dimensões representativas do objeto catalogado.

4.4 Metadata Object Description Schema (MODS)

A demanda de bibliotecas digitais e gestores de conteúdo implicou no advento do esquema MODS sob os auspícios do *Network Development and MARC Standards Office of the Library of Congress*. Sua criação e os testes iniciais se deram em 2002 (GUENTHER, 2003, p. 147). Embora a sintaxe seja semelhante ao MADS e MARCXML, sua finalidade consiste em descrever e representar tipos complexos de recursos de informação, o que inclui todos os campos e subcampos MARC21 destinados a codificar os pontos de acesso da descrição bem como estabelecer vínculos com remissivas e tratar adequadamente as formas variantes de nomes. Em complemento, Fachin (2011, p. 95) ressalta a interoperabilidade e compartilhamento como atributos dos dados codificados em MODS, ao mesmo tempo em que ressalta que esse esquema é mais completo que o DC e compatível com os dados bibliográficos. Na mesma perspectiva, e a exemplo de outros esquemas de metadados digitais para descrição bibliográfica, McCallum (2004, p. 84) aponta três pontos fortes do MODS:

- Incorpora as vantagens do ambiente XML;
- Oferece suporte especial à catalogação de recursos eletrônicos e é
- Amplamente compatível com o MARC21.

Observa-se que a sintaxe deste esquema possui escopo genérico, o que o habilita a representar todos os pontos de acesso que integram um recurso, ou objeto de informação catalogado. Uma autoridade de assunto e suas remissivas podem ser agrupadas de acordo com os atributos presentes na entidade a ser catalogada.

5 MODELAGEM CONCEITUAL DE METADADOS PARA O DOMÍNIO BIBLIOGRÁFICO

A representação conceitual, que envolve as entidades e atributos do domínio bibliográfico, tem sido fortemente influenciada pelo modelo entidade-relacionamento como paradigma capaz de fornecer uma visão de conjunto dos elementos que compõe as áreas da descrição. Tal modelo integra o conceito de modelagem conceitual, que segundo Fusco (2010, p. 99), expressa o relacionamento entre os dados além de fornecer uma perspectiva mais próxima sobre como os usuários visualizam tais relacionamentos.

Estes dados e seus relacionamentos, na condição de contributos diretos das etapas de mapeamento e codificação das autoridades aqui tratadas, são abordados segundo um olhar sobre a semântica das representações dos registros bibliográficos:

[...] a semântica de registros bibliográficos relaciona-se tanto ao [1] significado estabelecido entre cada um dos elementos representados e o documento representado, propondo-se um conteúdo semântico do registro bibliográfico, identificado como a semântica do registro bibliográfico em si, quanto ao [2] significado entre o elemento definidor e o valor dado para expressar este elemento, apresentado como subsemântica, no qual se estudam as unidades semânticas, e quanto ao [3] significado que pode ser estabelecido entre os vários registros de um catálogo, identificado como supersemântica. Em outras palavras: o registro bibliográfico de um documento só pode assumir o significado de representação daquele documento dado o sentido que pode ser estabelecido entre representado e representante, ou, entre obra e manifestação; pelo significado entre o elemento a ser representado e o seu valor (dado, por sua vez, pela sintaxe), e uma metasemântica, no sentido das relações de metalinguagem que podem ser estabelecidas entre diferentes registros. (ZAFALON, 2012, p. 14)

Entidade, por sua vez, é definida por Chen (1976) como uma 'coisa' que pode ser distintamente identificada, enquanto que um relacionamento é uma associação entre entidades. O modelo entidade-relacionamento surge a partir da proposta de Peter Chen que designou de "visão lógica de dados" o conjunto das entidades e relacionamentos em um modelo de domínio. Segundo Moreno (2007, p. 13), este modelo visa a captura da semântica dos dados e não a implementação.

O domínio bibliográfico na perspectiva dos FRBR, discutido no próximo capítulo, expressa o que o autor chama de "visão unificada de dados", uma vez que os atributos das entidades que integram o *framework* mantêm entre si uma relação coordenada.

Dentre as categorias que formam o núcleo da uniformidade dos catálogos e das coleções, Decourselle *et. al.* (2016, p. 254) discute quatro padrões que visam contemplar as entidades do modelo conceitual da IFLA: o padrão incremental, definido

como elemento adicional a itens existentes sem potencial para alterar seu conteúdo, tais como ilustrações e prefácios; padrão derivativo, baseado na modificação de um trabalho por outro, a exemplo das traduções e revisões de conteúdo; o padrão de agregação, descrito em termos da relação todo-parte a exemplo de autores e títulos coletivos, seus organizadores e autores; e o padrão de trabalhos complementares, que objetiva modelar o relacionamento entre trabalhos de importância análoga.

Estes padrões, além de assumirem especial relevância na designação dos responsáveis pelo conteúdo intelectual e nos atributos que o qualificam, são necessários na modelagem da relação que vincula autor ou organizador à expressão da criação artística ou intelectual, nos termos dos FRBR.

Tal relação é expressa pelas propriedades das sintaxes que integram os esquemas MADS e MODS, quando a intenção consiste em tratar dados de autoridade e assunto ou o METS, quando a finalidade é a descrição do objeto ou a manifestação da expressão. Em esquemas mais complexos como o vocabulário RDF, é comum atribuir vínculos e ligações concatenados a um identificador específico, o URI.

Como elemento integrante da representação conceitual de um acervo bibliográfico, o identificador de recursos uniforme ou URI permite, nas palavras de Ferreira e Santos (2013, p. 18), nomear e identificar recursos abstratos, sejam eles processáveis ou não por computador. Os mesmos autores ainda apontam detalhes adicionais fornecidos pelas URI's, tais como dados de publicação de livros e disponibilidades para empréstimos. Um URI inicia com o protocolo *http* no formato compatível com a web. É frequente encontrar URI integrado a vocabulários formais como ontologias de domínio e o modelo para descrição de recurso ou RDF, uma vez que estes esquemas, conforme lembram Catarino e Souza (2012, p. 79), fazem uso dos dados lincados, ou *linked data* em inglês, cuja finalidade consiste em relacionar coleções de dados.

5.1 Functional Requirements for Bibliographic Records (FRBR)

As discussões que deram origem aos Requisitos Funcionais para Registros Bibliográficos ou FRBR surgiram de encontros e debates em torno das regras da catalogação em nível internacional, ratificados pela Conferência Internacional sobre Princípios de Catalogação realizada em Paris no ano de 1961. E, posteriormente, pela Reunião Internacional de Especialistas em Catalogação (RIEC) em 1969 (ALVES, 2010, p. 83). Neste aspecto, Bezerra (2013, p. 92) justifica os FRBR como "marco

fundamental da contemporaneidade", no tocante a adequação de formatos e padrões ao modelo conceitual proposto e na revisão do ISBD.

As prescrições normativas para escolha dos pontos de acesso e, das entradas autorizadas, passaram a considerar o provimento de um modelo conceitual claro e definido apto a relacionar dados codificados em registros bibliográficos em nível lógico que integra atributos e entidades; além de recomendar um nível básico de funcionalidade para os registros criados por agências bibliográficas nacionais. Segundo Tosaka e Park (2013, p. 653), a expansão dos padrões, formatos e procedimentos afetos à catalogação tem se caracterizado pela reconfiguração do AACR2 para o RDA, incorporando os modelos FRBR/FRAD.

No relatório final *final report* do FRBR publicado em 2008, encontramos outro grande objetivo discutido pela IFLA: a inclusão das chamadas *user tasks*, ou atividades do usuário, na medida em que este assume diversos papéis na criação, organização, tratamento e acesso à produção bibliográfica.

No entanto, a finalidade dos requisitos funcionais é, nas palavras de Fusco (2010, p. 19) operar como "modelo de referência" e no "nível conceitual". Não existe na proposta da IFLA, a operacionalização do *framework* ao nível de implementação em um sistema de banco de dados ou desenvolvimento de aplicações.

Quanto à sua estrutura, os FRBR são compostos por entidades, tratadas como "[...] objetos-chave que interessam ao usuário do registro bibliográfico" (SILVEIRA; TÁLAMO, 2009, p. 110). As entidades, por sua vez, são categorizadas em três grupos:

Grupo 1: Os produtos do trabalho intelectual que contempla a obra, expressão, manifestação e item;

Grupo 2: Os responsáveis pelo conteúdo das entidades do primeiro grupo, que envolve pessoa e entidade coletiva; e

Grupo 3: Entidades que representam os assuntos de uma obra: conceito, objeto, evento e lugar.

Outro componente integrante da estrutura do modelo conceitual são os atributos, definidos como meios pelos quais os usuários formulam consultas e interpretam respostas quando da busca de uma entidade em particular. (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2008, p. 31). Uma entidade pode possuir atributos que lhes sejam inerentes e atributos externos que lhes sejam associados. Estes últimos devem fazer referência às fontes externas que atuam como instâncias da entidade, a exemplo do

equipamento necessário para reprodução de uma mídia. O quadro 5 do capítulo 5.2 lista um conjunto de atributos considerados inerentes às entidades correspondentes.

Cabe destacar o uso recorrente do termo inglês *frbrisation*, sempre que existe ação de mapear e extrair das entidades seus atributos e relacionamentos no modelo conceitual do FRBR a partir da análise de um registro bibliográfico existente.

Para Aalberg e Zumer (2012, p. 854), o foco atual da *frbrisation* traz limitações que demandam uma abordagem mais sistemática para interpretar os dados de um registro bibliográfico. Tais limitações incluem: manifestações em uma única expressão, manifestações em múltiplas expressões, manifestações e trabalhos relacionando agentes individuais e coletivos, relacionamento entre assuntos dentre outros.

É neste contexto que a IFLA, ao propor os FRBR, ressalta a relevância do usuário e dos registros bibliográficos em termos de conteúdo e estrutura IFLA (2008, p.4), como forma de priorizar as variáveis que possam vir a interferir nos custos da catalogação e na duplicidade do exercício catalográfico.

5.2 Functional Requirements for Authority Records (FRAD)

Os requisitos funcionais para registros de autoridade (FRAD), tem sua origem com a criação de um grupo de trabalho estabelecido em abril de 1999, denominado *Functional Requirements and Numbering of Authority Records* (FRANAR), sob a responsabilidade da IFLA *Division of Bibliographic Control* e a IFLA *Universal Bibliographic Control and International MARC Programme* (UBCIM). Segundo Maimone, Silveira e Tálamo (2011, p. 32), ao considerarem em conjunto os modelos FRBR, FRAD e FRSAD, destacam que seus elementos constitutivos foram desenvolvidos como uma visão geral do universo bibliográfico, relacionando os dados contidos nos registros bibliográficos de acordo com as necessidades dos usuários.

Os requisitos para autoridades são tratados pela IFLA como uma "extensão" e "expansão" do modelo que lhe deu origem (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2008, p. 2). Subsidiariamente, pois, as entidades abordadas pelos FRAD têm como origem os FRBR e o *Guidelines for Authority Records and References* (GARR) que, por sua vez, integra o escopo do FRAD à proposta de modelagem conceitual do domínio bibliográfico.

Para cada entidade mencionada no relatório da IFLA, existem inúmeras possibilidades de vínculos, perpassando pelos 3 grupos e em ampla consonância com o

ambiente *web*, em particular, os esquemas de metadados aplicados à descrição bibliográfica tratados nos capítulos seguintes desta dissertação.

Embora o enfoque da IFLA seja o estabelecimento de um modelo conceitual concernente ao universo bibliográfico, destacou-se para esta pesquisa aqueles atributos presentes nos grupos 2 e 3, uma vez que é nestes grupos aonde encontramos as entidades e atributos de pessoas e assuntos.

No âmbito dos registros de autoridade, os relacionamentos são efetivados pelos atributos que integram a entidade a ser representada. O quadro 4 representa, em linhas gerais, o panorama de relacionamento entre atributos e as entidades relativas aos dados de autoridade abordados pelo FRAD.

Quadro 4: Relacionamentos entre nomes pessoais, de família, entidade e obra

Tipo de Entidade	Amostra de relacionamentos
Nome Pessoal	<ul style="list-style-type: none"> • Relação de nome anterior • Relação com nome atual • Relacionamento com forma linguística alternativa • Outras variações de nomes
Nome de Família	<ul style="list-style-type: none"> • Relacionamento com forma linguística alternativa
Nome da Entidade	<ul style="list-style-type: none"> • Relação com nome estendido • Relação com acrônimos / iniciais / abreviações • Relacionamento com forma linguística alternativa • Outras variações de nomes
Nome da Obra	<ul style="list-style-type: none"> • Relacionamento com forma linguística alternativa • Relacionamento com nome convencional • Outras variações de nomes

Fonte: adaptado da INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2013)

Em seu relatório, o FRAD lista as entidades úteis ao modelo que complementam a lista do quadro 4, quais sejam: expressão, manifestação, item, conceito, objeto, eventos, lugar, nome, identificador, ponto de acesso controlado, regras e agências.

O caráter não exaustivo e não prescritivo é ressaltado pelos FRAD, uma vez que tais entidades, e os atributos que lhes são correlatos, poderão ser utilizados na

medida e extensão em que forem demandados pelos sistemas de informação, em consonância às necessidades dos usuários do catálogo.

Uma entidade comporta atributos que possuem características que auxiliam na modelagem conceitual do domínio. O FRAD destaca, por exemplo, a posição dos atributos ao nível lógico do modelo, bem como os elementos que atuam como qualificadores das autoridades a que se conectam. Os atributos estão também sujeitos a modificações ao longo do tempo, como mudanças de nomes, títulos de nobreza e cargos dentre várias modificações possíveis.

Particularmente útil aos propósitos desta pesquisa, estão aqueles atributos que melhor representam uma entrada autorizada de nome, e que, posteriormente, serão mapeados para os esquemas de metadados em lide. O quadro 5, agrega os atributos do tipo pessoa, que serão adotados no mapeamento dos capítulos 6.1 e 6.2 e que, segundo a INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2013, p. 17), representam complementos aos já existentes no FRBR.

Quadro5: Atributos de autoridade de nome pessoal.

Datas associadas	Datas associadas à pessoa, o que pode contemplar dia/mês/ano do nascimento e morte ou século de atividade.
Título da pessoa	Título de chefes de estado (presidentes, governadores) e de nobreza (reis, duques e príncipes).
Gênero	Gênero pelo qual a pessoa é identificada (masculino, feminino, desconhecido e outros)
País	País com o qual a pessoa é identificada
Idioma	Idioma utilizado pela pessoa na criação de uma expressão para publicação, transmissão etc.
Afiliação	O grupo ao qual a pessoa é afiliada ou foi afiliada na condição de membro(a), empregado(a),por identidade cultural etc.

Fonte: adaptado de INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2013, p. 17)

Os atributos arrolados no quadro 5 foram extraídos pelos FRAD do FRBR, GARR, UNIMARC *Manual – Authorities Format* e o *Mandatory Data Elements for Internationally Shared Resource Authority Records* (MLAR). A exaustividade

necessária se ajusta ao paradigma entidade-relacionamento na medida em que contempla os componentes do modelo representativo dos dados de autoridade.

Os relacionamentos dos pontos de acesso da descrição propiciam ao usuário percorrer os registros de um catálogo Silveira e Tálamo (2009, p. 114), cumprindo sua função de estabelecer vínculos entre entidades e atributos. Neste aspecto, passamos da análise dos requisitos de autoridade FRAD para os requisitos de autoridade de assuntos FRSAD, mantendo a perspectiva do paradigma entidade-relacionamento.

5.3 Functional Requirements for Subject Authority Data (FRSAD)

Os FRSAD nascem da iniciativa da IFLA de extrair do grupo 3 as entidades que integram o relacionamento entre os recursos bibliográficos e os assuntos. Segundo o relatório final dos FRSAD, é aplicado às entidades do grupo 3 o modelo conceitual do FRBR, que aborda o conceito como (noção abstrata ou idéia), um objeto (coisa material), evento (ação ou ocorrência) e lugar (uma localização). (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2010, p. 14)

O modelo conceitual para os dados de autoridade de assunto é norteado pelas tarefas do usuário, ou *user task*, discutidas no capítulo 5.1, que no âmbito dos FRSAD são traduzidas conforme quadro 6:

Quadro 6: Tarefas gerais do usuário no escopo do FRBR

Encontrar	Um ou mais assuntos e / ou suas denominações, que correspondem aos critérios estabelecidos pelo usuário, utilizando atributos e relacionamentos.
Identificar	Um assunto e / ou sua denominação com base em seus atributos ou relações (isto é, para distinguir entre dois ou mais assuntos ou denominações com características semelhantes e para confirmar que o assunto é apropriado ou que a denominação foi encontrada);
Selecionar	Um assunto e / ou sua denominação apropriada às necessidades do usuário (isto é, escolher ou rejeitar com base nos requisitos e necessidades do usuário);
Explorar	As relações entre os sujeitos e / ou as suas denominações (por exemplo, para explorar relações a fim de compreender a estrutura de um domínio sujeito e a sua terminologia)

Fonte: INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2013).

O conjunto de tarefas no quadro 6 expressa a relação entre os usuários e o universo bibliográfico, em particular, no domínio de assunto, o que envolve criadores e gestores de metadados, usuários finais, profissionais da informação dentre outros.

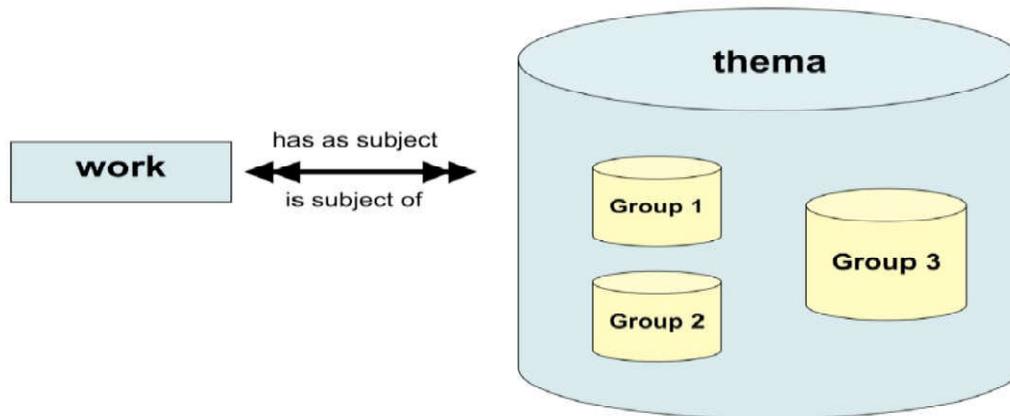
Para melhor formalizar o modelo conceitual dos FRSAD, o grupo de trabalho atribuiu aos requisitos duas designações latinas. A primeira chamada *thema* que se refere a qualquer entidade usada como assunto de uma obra, a segunda, *nomen* que corresponde a qualquer convenção simbólica ou signos em que um *thema* é conhecido. A opção pelo latim para designar as duas entidades dos FRSAD é justificada pela IFLA por não possuírem significados pré-existentes no contexto ocidental, por serem culturalmente neutras e não requererem tradução. (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2010, p. 16)

Os atributos que integram os relacionamentos entre as entidades *thema* e *nomen* seguem normas de restrição a depender do contexto em que são empregadas, com possibilidade de relacionamento de um para um, um para muitos e/ou muitos para muitos conforme veremos no esquema a seguir:

Devido à entidade *thema* ser definida de uma forma muito abstrata e geral, além de abranger tudo o que pode ser assunto de uma obra, os atributos de um *thema* são dependentes do contexto de sua implementação e irão variar, consideravelmente, em diferentes situações de uso. Por isso, há somente dois atributos aplicáveis em todos os casos: tipo de *thema* e nota de escopo. Todos os outros atributos dependem do contexto específico de aplicação a que um determinado *thema* pertence. (MELO, 2013, p. 59)

A exaustividade da representação de um *thema* na relação do tipo muitos-para-muitos é representada pelos FRSAD e envolve os três grupos dos requisitos funcionais, o que o inclui como uma classe geral na categoria de assuntos conforme figura 9:

Figura 9: Relacionamento muitos para muitos entre *work ethema*



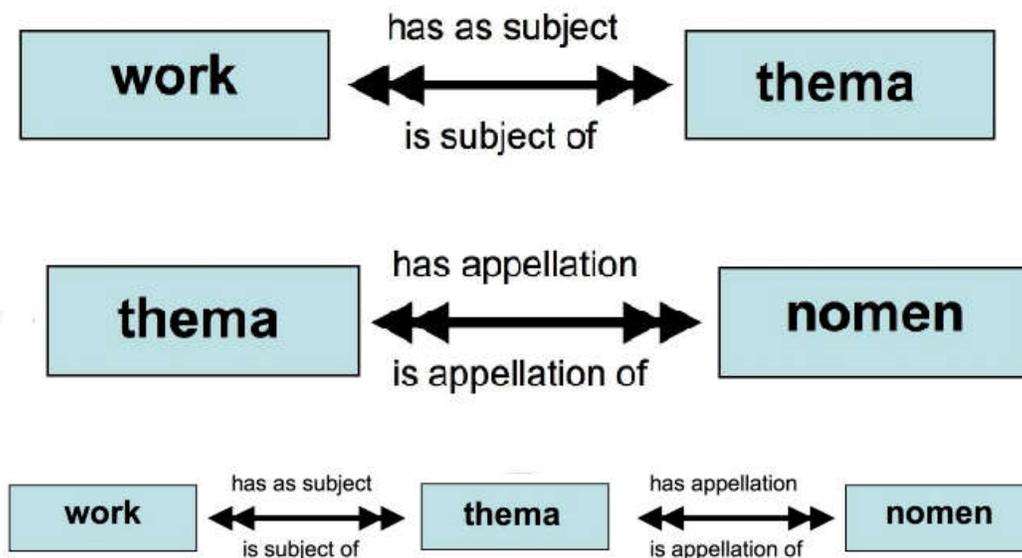
Fonte: INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2010)

Os diversos atributos de um recurso de informação e as entidades que eles podem representar implicam um modelo capaz de esclarecer a relação dinâmica a ser estabelecida entre um *work* e todas as entidades dos 3 grupos dos FRBR.

Esta compreensão é necessária para que a interdependência entre os diversos atributos de uma entidade não se confunda com seu isolamento no domínio bibliográfico. Razão pela qual o diagrama da figura 9 esclarece por meio das setas a relação necessária entre os atributos de *work* e *thema*.

Isto posto, e como partição do relacionamento genérico representado pela figura 9, temos que a mesma paridade nos relacionamentos existe entre *thema* e *nomen*, uma vez que um *thema* pode estar vinculado a um ou mais *nomen* assim como um *nomen* pode estar associado a um ou mais *thema*. No entanto, são as necessidades dos usuários que irão determinar quais atributos podem ser considerados úteis ao universo que se deseja representar:

Figura 10: Relacionamentos possíveis entre *work*, *thema* e *nomen*.



Fonte: INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2010)

Em sua pesquisa, Melo (2013, p.60) fez uma análise detalhada sobre o modelo apresentado, afirmando se tratar de uma listagem não exaustiva por reconhecer que nem todos os atributos listados são aplicáveis a todos os *nomens*. Assim, os atributos tipo: esquema, fonte, representação, linguagem, escrita e conversão de escrita, são descritos como aplicáveis em contextos específicos de sistemas de organização do conhecimento.

Ao definir as estratégias de conversão dos dados de autoridades de nome e assunto, considera-se a interdependência lógico-funcional do modelo entidade-relacionamento e dos esquemas de metadados no papel de descrever e representar sintática e semanticamente o universo bibliográfico.

Ao mesmo tempo, os esquemas de metadados MARCXML, MADS e MODS propiciam a expansão de notações destinadas a tratar ambiguidades, polissemias e demais inconsistências na descrição bibliográfica, conforme abordado nos capítulos seguintes destinados à etapa empírica.

6 ESTUDO EMPÍRICO: MAPEAMENTO E CONVERSÃO DOS PONTOS DE ACESSO DE AUTORIDADES DE NOME E ASSUNTO

Nos capítulos 6.1 e 6.2 foram implementadas as estratégias de mapeamento e *crosswalking* nas categorias de autoridade selecionadas para esta pesquisa. Necessário se faz reiterar que a abordagem dos grupos 2 e 3 dos FRBR em correlação direta com os metadados em estudo e como caminho para atingir os objetivos propostos, visa ir ao encontro dos paradigmas da modelagem conceitual, discutidos no capítulo 5 e cuja etapa se faz imprescindível na definição das estratégias desenvolvidas a seguir.

A estrutura semântica de um registro bibliográfico e das autoridades que o integra, por sua vez, depende de atributos sem os quais a representação dos pontos de acesso e uso dos catálogos resultam comprometidos. Tais atributos, como os de família, possibilitam a identificação unívoca de um dado nome pessoal.

Os capítulos 6.1 a 6.5, ao delinearem as etapas da conversão dos metadados de autoridades, tratam, no seu decorrer, dos atributos e das diversas modalidades de seu uso no catálogo por meio dos esquemas MARCXML, MODS e MADS.

Estas tarefas abordam diretamente as categorias de análise selecionadas para esta pesquisa e constantes das diretrizes metodológicas, quais sejam as autoridades de nomes e de assuntos. Além da perspectiva estruturalista na composição das hierarquias, associações e vínculos de entidades e notações do domínio bibliográfico.

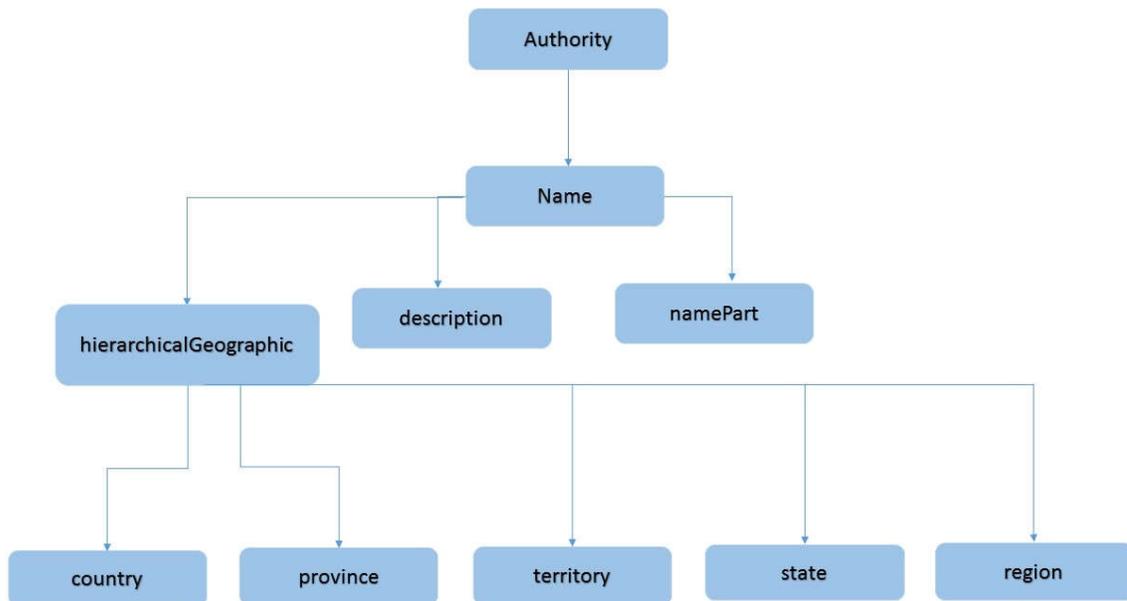
6.1 Mapeamento das autoridades de nomes

Nos catálogos de autoridades de nomes, são encontradas categorias diferentes de entradas autorizadas que merecem atenção com relação às suas especificidades. Dentre essas entradas, as mais comuns e observadas nos mapeamentos realizados são as de pessoa, entidade, eventos e nome geográfico.

A hierarquia inerente ao esquema XML permite situar do genérico ao específico os *namespaces* presentes em MADS, o que facilita a posterior inserção das propriedades com finalidade relacional.

O ponto de partida do mapeamento é representado na figura 11, elaborada a partir do esquema MADS que estabelece a hierarquia dos metadados que o integra, em que o atributo-pai sinaliza para dados de autoridade de nome pessoal.

Figura 11: Hierarquia de elementos de autoridade de nome em MADS.



Fonte: Elaborado pelo autor.

Os elementos de autoridade *Name* na figura 1 são desdobrados pelos FRAD em atributos de entidades destinados a assumir diversas modalidades de manifestação. Tais atributos integram o modelo relacional da entidade representada e seus vínculos têm o condão de qualificar o relacionamento. Os atributos-filho *description* e *namePart* possuem campos variáveis cuja finalidade consiste em inserir informações complementares, no caso de *description*, e propriedades atribuídas no caso de *namePart*. Segundo a terminologia adotada pelo MADS, tais atributos recebem a designação de subelementos. Importa destacar que o subelemento *namePart* aparece vinculado a atributos qualificadores, dentre eles *date* e *family* que, por sua vez, são encontrados de forma mais recorrente nos cabeçalhos de autoridade de nomes. No esquema MADS, as formas variantes de nomes aparecem representadas pela propriedade *variant type* que figura como elemento de nível superior atuando como container em relação às propriedades *Name* e *namePart*. No capítulo destinado à codificação em MADS esta propriedade será melhor explorada em termos de cotejamento com os campos no formato MARC correspondentes.

No quadro 7, foram mapeados os atributos concernentes a dados de autoridade de nome, como uma especialização das diferentes facetas que esta categoria pode assumir no catálogo de autoridades.

Quadro 7: Mapeamento dos atributos de Nome segundo FRAD.

Atributos de Nome	Tipos de Nomes				
	Nomes pessoais	Entidades	Nomes de famílias	Títulos de trabalhos	Manifestações
	<i>NameString</i>				
	Uma sequência de caracteres alfabéticos e/ou numéricos que representam o nome de uma entidade.				
	Escopo de uso				
	Forma de trabalho associada a um nome particular.				
	Datas de uso				
	Datas associadas com o uso de um nome particular estabelecido para uma pessoa, entidade ou família.				
	Idioma				
	A língua na qual o nome é expresso.				

Fonte: Adaptado de INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2013, p. 24)

Os atributos do quadro 7, quando inseridos em catálogos de autoridades, ampliam as instâncias de um nome pessoal por meio da geração de pontos de acesso adicionais. Ao mesmo tempo, promove fator de relevância quando se planeja diferenciar autores pessoais, ou mesmo nomes corporativos semelhantes.

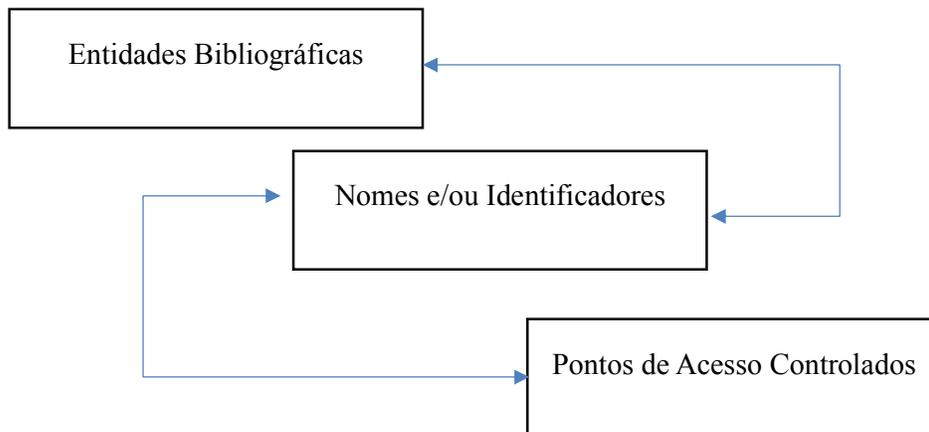
Embora o mapeamento realizado para esta pesquisa trate de nomes pessoais, observa-se em Tipos de Nomes no quadro 7, que os itens listados na sequência como: escopo de uso, datas e idiomas atuam em consonância com a busca de maior consistência e especificidade na individualização de uma dada autoridade independente de qual categoria de entidade pertença.

Ao adotar um modelo de cunho relacional, a compreensão do universo bibliográfico abandona uma perspectiva linear e pragmática de relação unidirecional entre autor e obra e passa a vislumbrar as entidades, atributos e suas relações como um novo fator estratégico de redução de perdas de informação.

A estratégia de mapeamento de uma autoridade de Nome Pessoal, segundo modelo entidade-relacionamento adotado pelos FRBR e discutido no capítulo 5, condiciona o uso de qualificadores de função atribuídos ao autor, quais sejam: ilustrador, prefaciador, tradutor dentre outros, bem como identificadores que relacionam um autor a determinada obra ou quais assuntos são abordados por determinado autor

segundo atributos como os listados na figura 11. Um diagrama que sintetiza e orienta esse relacionamento do tipo *many-to-many* ou muitos para muitos é ilustrado pelos FRAD na figura 12:

Figura 12: Fundamentos para o Modelo Conceitual.



Fonte: INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (2013, p. 4)

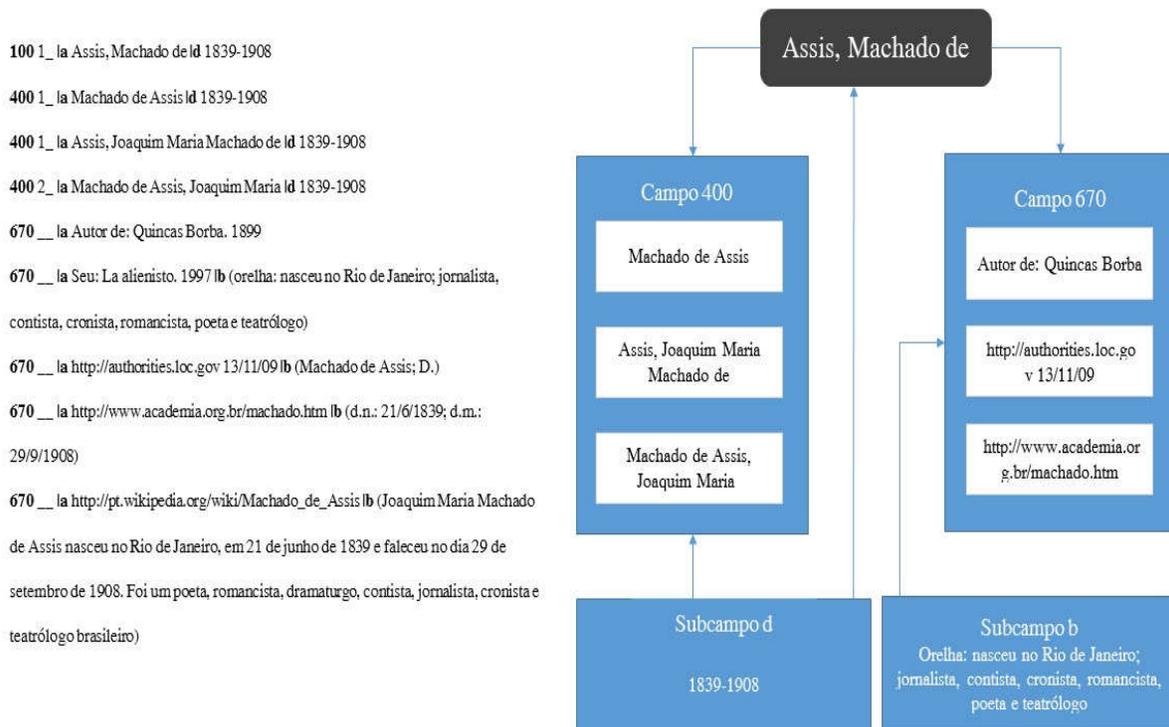
O esquema que vincula as Entidades Bibliográficas aos Pontos de Acesso Controlados, conforme representado na figura 12, constitui importante diferencial no âmbito dos esquemas XML, uma vez que sua estrutura comporta o uso de identificadores.

Existem, no entanto, diversas modalidades para definição dos identificadores destinados a representar adequadamente as entradas autorizadas para concretização do modelo da figura 12. Alguns destes identificadores são demonstrados no capítulo 6.4 da página 70, e não se confundem com o item dos FRAD denominado "Atributos de um Identificador", uma vez que, neste caso, trata-se de identificação das obras por meio dos ISBN e ISSN correspondentes. Verifica-se que o controle e a padronização dos dados na descrição de pessoas e entidades se consolidam como fator de confiabilidade na modelagem conceitual de catálogos de autoridades. Tais dados demandam levantamentos das histórias das famílias em sua origem bem como das árvores familiares, como requisito de redução de ambigüidades e promoção de maior acurácia na relação dos nomes às suas entidades, manifestações e expressões na terminologia dos FRBR.

O mapeamento das autoridades de Nome Pessoal considerou sua ocorrência segundo a função exercida pelas referidas autoridades nos documentos analisados. Trata-se de obras de autoria individual, até três autores ou mais de três autores, que figuram como responsáveis pelo conteúdo intelectual da obra e de organizador, comumente atribuído a coletâneas e anais de eventos e as entradas secundárias.

Os campos e subcampos correspondentes no formato MARC tratados no capítulo 3.1.1.1 da página 29, indicam a sintaxe adotada para cada responsabilidade de autoria mencionada no parágrafo anterior, bem como a forma que devem figurar no cabeçalho com seus respectivos indicadores. A figura 13 apresenta o mapeamento para o modelo entidade-relacionamento dos campos e subcampos de uma autoridade de nome pessoal:

Figura 13: Mapeamento de dados atribuídos a Autor Pessoal.



Fonte: Elaborado pelo autor.

Os atributos de nome pessoal representados na figura 13 são pontos de acesso para o cabeçalho autorizado. Na coluna da esquerda, foi incluído o campo 100 para

designar autoria principal representada, seguida de três ocorrências do campo 400 destinadas a relacionar remissivas ou formas variantes para a autoridade do campo 100.

Os campos 670 são caracterizados como fonte positiva de dados, ou as fontes de onde os nomes arrolados anteriormente foram retirados, bem como outras informações conforme veremos adiante. Tais fontes são comumente extraídas de listas de cabeçalhos de assuntos, tesouros e bases de autoridades, com a necessária observância de que, na determinação do nome, deve-se dar preferência a sua forma mais conhecida, motivo pelo qual na figura 13, Machado de Assis foi vinculado ao campo 100.

O mapeamento para autoridade pessoal que figura como organizador de coletâneas está representado no quadro 8, que estabelece os principais pontos de acesso definidos pelo AACR2 para este tipo de documento. Conforme mencionado anteriormente, o papel atribuído aos nomes pessoais atende às normas prescritas pelos FRAD para dados de autoridade, que no formato MARC é caracterizado pelo termo de relação e representado pela letra "e" como subcampo do campo 700 - Entrada Secundária de Nome Pessoal.

Quadro 8: Registro MARC para uma coletânea.

Remissivas para autoria pessoal			
020		9788563478023 (broch.)	
245	0 0	\$a Ações institucionais de avaliação e disseminação de tecnologias educacionais / \$c organizadores: Ricardo Azambuja Silveira, Raymundo Carlos Machado Ferreira Filho	
260		\$a Porto Alegre: \$b JSM Comunicação, \$c 2011	
300		\$a 202p. : \$bil. ; \$c 21cm.	
504		\$a Inclui bibliografia.	
700	1	\$a Silveira, Ricardo Azambuja	\$e Org.
700	1	\$a Ferreira Filho, Raymundo Carlos Machado	\$e Org.

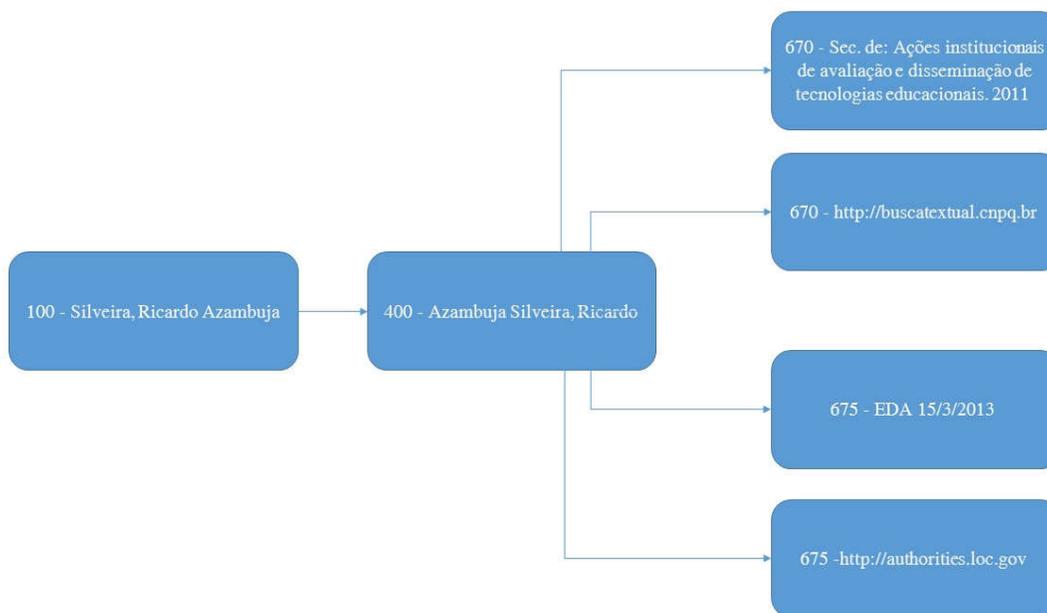
Fonte: Elaborado pelo autor.

Como alternativa para modelar a relação entre autor pessoal e seu papel na obra, o formato MARC Bibliográfico possui o campo 700 destinado ao registro das entradas secundárias para autor pessoal. Seu uso é requisito em coletâneas cuja entrada da obra se dá pelo título, bem como naqueles casos em que a folha de rosto, fonte principal de informação para o catalogador, não cita o autor pessoal como responsável intelectual pelo seu conteúdo. No entanto, o mesmo se faz presente como segundo autor, tradutor, revisor, organizador dentre outras atribuições. Aqui, mais uma vez, temos o uso

do subcampo como atributo qualificador do papel desempenhado pela autoridade de nome pessoal. Ao ser adotado o ponto de acesso secundário, ou campo 700, o catalogador especifica no subcampo \$e sua função no todo ou na parte da obra sob descrição.

O diagrama a seguir representa, além da forma variante de nome, os vínculos do autor com as produções das quais fez parte. São arroladas aqui, as atribuições nas quais o autor figura sob diversas formas além de dados biográficos e filiações institucionais. Exemplo do aspecto heterogêneo representado pelo campo 670.

Figura 14: Mapeamento de um registro de autoridade de Nome Pessoal



Fonte: Elaborado pelo autor.

Opcionalmente, poder-se-ia criar uma remissiva neste diagrama para a citada coletânea bem como a outras obras de sua autoria. Além dos campos 670 e 675 tratados anteriormente, o nome do autor ou organizador de uma obra pode aparecer como assunto; neste caso, o MARC atribui o campo 600 cujos mapeamentos com os demais campos serão abordados na subseção 6.2 seguinte.

O relacionamento entre entidades de Nome Pessoal, segundo terminologia do FRBR, pode tratar de nomes não controlados, ou seja, nomes que não constam em listas de autoridades ou possuem formas dissonantes daquelas estabelecidas no AACR2. O

formato MARC atribui o campo 720 para esses casos, e assim como no campo 700 tratado no parágrafo anterior, o termo de relação com o documento codifica-se com o subcampo \$e.

No capítulo seguinte, o mapeamento para assuntos prepara os campos MARC destinados a codificação nos esquemas XML segundo a profundidade e extensão dos termos tópicos tratados.

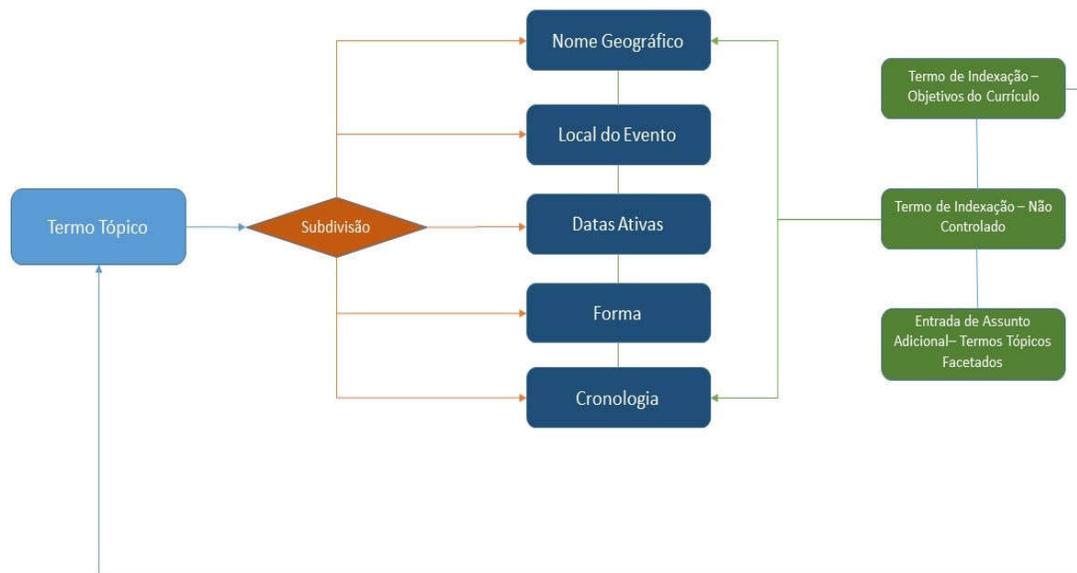
6.2 Mapeamento das autoridades de assuntos

No capítulo 5.3 da página 52 foram apresentados e discutidos os requisitos para os dados de autoridade de assuntos, representados no contexto dos FRBR pelo acrônimo FRSAD. Dos tipos de relações estabelecidas entre *work*, *thema* e *nomen* importa ao mapeamento das entidades do grupo 3 dos FRBR identificar campos e subcampos no formato MARC aptos a modelar e persistir dados de identificação com obras e assuntos.

No que concerne a estes últimos, o relatório final dos FRSAD compara classes e instâncias a categorias universais e particulares de um domínio de assunto representado em bases de autoridades, semelhante ao estabelecimento de categorias hierárquicas frequentemente usual nos sistemas de classificação de assuntos.

A etapa preliminar do mapeamento para autoridades de assuntos, a exemplo do capítulo anterior, consistiu em *crosswalking* do esquema MODS para um diagrama entidade-relacionamento. O resultado do cotejamento entre os campos e subcampos do formato MARC com as propriedades e atributos das *tags* correspondentes nos esquemas de metadados, MARCXML, MADS e MODS, abordou a modelagem para os dados de autoridades de nomes. Dentre os subcampos para autoridade de assunto, por sua vez, são destacados: \$a termo tópico, \$c locais de eventos, \$v subdivisão de forma, \$y subdivisão cronológica e \$z subdivisão geográfica. As entradas catalogadas nestas subdivisões podem ser remissivas para termos tópicos se assim a obra o permitir. É o caso dos mapas, convenções cartográficas, escalas e atlas geográficos, cujos nomes de cidades, estados e países podem tanto estar representados na forma de subdivisão como na forma de termo tópico ou campo 651 do formato MARC. Uma biografia incluirá o campo 600 para designar o nome pessoal do biografado que aqui é assunto da obra, ao mesmo tempo em que também aparece no campo 100 quando seu papel figurar como autor pessoal.

Figura 15: Termo Tópico e remissivas aplicáveis a subdivisões.



Fonte: Elaborado pelo autor.

Na figura 15, representou-se a relação entre o Termo Tópico ou assunto, constante do grupo 3 dos FRBR, com algumas de suas subdivisões ao centro relacionadas entre si por uma linha vertical. Na relação à direita, temos alguns campos que atuam na especificação do assunto tratado no Termo Tópico, quais sejam: campo 658 – Objetivos do currículo, 653 – Termo de Indexação não Controlado e o 654 – Entrada de Assunto Adicional Termos Tópicos Facetados. O relacionamento entre as entidades inclui parâmetros auxiliares para a modelagem conceitual, tais parâmetros consistem nos pontos de acesso, aqui representados pelos campos do formato MARC, e os identificadores que constituem os fundamentos da modelagem. Já os subcampos associados serão objeto do quadro 9, e servem de parâmetro para a posterior codificação nos esquemas de metadados dos capítulos seguintes.

Os identificadores podem auxiliar na individualização de uma obra por meio do ISBN ou ISSN correspondentes, ou podem ser representados por sequências de caracteres alfanuméricos gerados dinamicamente por um banco de dados. Ressalta-se que os sistemas de automação de bibliotecas utilizam esses identificadores como elo de ligação entre a obra, autoridade e assunto na forma de mapeamento objeto-relacional.

No paradigma entidade-relacionamento aqui adotado, os esquemas de metadados baseados em XML e analisados nos capítulos seguintes, incorporam esses identificadores com vistas a estabelecer expressividade semântica entre autoria e assunto.

No mapeamento realizado a hierarquia está presente, ainda que de forma implícita, na própria estrutura dos campos de autoridades de assuntos. Tais hierarquias são identificadas por meio dos subcampos como os mencionados nos parágrafos anteriores.

Contudo, para melhor representar os termos tópicos referentes a assuntos geográficos, o formato MARC Bibliográfico possui o campo 662 – Entrada Adicional de Assunto – Nome de Lugar Hierárquico. A disposição dos subcampos a ele associados possuem uma ordem hierárquica explícita que parte do geral para o particular.

662 \$a Brasil \$b Ceará \$c Fortaleza.

A ordem hierárquica implícita no termo tópico representado pelo campo 650, por sua vez, não possui uma estrutura de subcampos para hierarquizar tipos como no exemplo do campo 662. Sua disposição é estabelecida pela abrangência do termo que segue do geral ao específico:

- 650 \$a Delinquência.
- 650 \$a Delinquência Juvenil.
- 650 \$a Delinquência Juvenil \$x Aspectos Psicológicos.
- 650 \$a Delinquência Juvenil \$x Prevenção.

Observa-se, no entanto, que os campos presentes no quadro 10 podem criar instâncias de um assunto tópico ou, ainda, ampliar a hierarquia das facetas conceituais que um termo pode assumir.

A partir da relação acima exemplificada, o campo 653 está apto a ser aplicado no desdobramento de assuntos representados por termos que não constam em bases de autoridades, sem embargo da coerência na disposição hierárquica da terminologia aqui tratada.

Neste momento, é possível visualizar a amplitude do escopo descritivo do formato MARC, mesmo quando existe a necessidade de representar a multidimensionalidade de um assunto tópico presente em forma de hierarquias implícitas, como a relação representada pelo campo 650.

Quadro9: Relação de campos e subcampos para definição de hierarquia de assuntos.

Campo	Subcampo	Definição	Exemplo
658 – Objetivos do Currículo	\$a	Principal objetivo do currículo	Objetivos da leitura (ficção)
	\$b	Objetivo de currículo subordinado	Compreensão da linguagem, temas e enredos.
	\$c	Código do currículo	NRPO2-1991
	\$d	Fator de correlação	Alta correlação
653 – Termo de Indexação não Controlado	\$a	Termo não Controlado	Dublin
654 – Entrada de Assunto Adicional	\$a	Termo Foco	Design de interiores
	\$b	Termo <i>Non-Focus</i>	Linhas contemporâneas
	\$c	Designação Faceta-Hierarquia	z
	\$e	Termo Relacionado	Decoração
	\$v	Subdivisão de Forma	2 v.
	\$y	Subdivisão Cronológica	1995
	\$z	Subdivisão Geográfica	São Paulo, SP

Fonte: Adaptado de LIBRARY OF CONGRESS (2017).

Ordenar subcampos de assuntos e dados relacionados a autoridade de nome pessoal em formatos legíveis por humanos e máquinas, permite ampliar o escopo das subdivisões hierárquicas e estabelecer vínculos entre atributos apropriados às entradas.

A DTD tratada no capítulo 3.2.1 da página 34 aplica instruções de validação a um documento MARCXML, o que reflete na representação dos dados de autoridade e na subsequente codificação das hierarquias de tipos de dados conforme apresentado nos esquemas XML tratados a seguir.

6.3 Codificação em MARCXML

As marcações de um esquema MARCXML são codificadas a partir de um DTD padrão, cuja aplicabilidade aos registros bibliográficos encontra-se detalhado no capítulo 3.2.1 da página 34. Segundo Flaminio (2006, p. 121), o formato USMARC figura como DTD para o MARCXML, acrescido de indicadores, campos e subcampos que seguem a formatação da ISO 2709.

Ao efetuar a conversão dos metadados em lide para interfaces do usuário são utilizadas tecnologias como XSLT e as *Cascading Style Sheets* (CSS), cuja formatação para leitura em navegadores é efetivada com o auxílio de analisadores sintáticos ou *parsers* que traduz para a linguagem HTML os *namespaces* do documento MARCXML com seus respectivos atributos e propriedades descritivas.

Os campos destinados ao controle do registro, tais como o Líder e Campos de Controle, são também representados no prólogo do esquema MARCXML, com *tags* identificadas para cada campo. Esta característica do esquema tem como premissa reproduzir o conteúdo do registro sem perda de informação.

A partir do MARCXML, é possível selecionar e extrair os pontos de acesso da descrição das autoridades que se deseja representar no catálogo. A codificação de um registro bibliográfico em MARC21, ilustrado na figura 16, relaciona as principais áreas da descrição, das quais codificou-se em MARCXML somente os respectivos campos 100 e 650, Autoridade de Nome Pessoal e Assunto Tópico respectivamente.

Figura 16: Aplicação do MARCXML na codificação de dados de autoridade de Nome Pessoal e Assunto Tópico.

005		20040614103959.0	
020		\$a 8586584053: \$c (broch.)	
082	0 4	\$a 155.412	<datafield tag="100" ind1="1" ind2=" " >
090		\$a 155.412 \$b S171a	<subfield code="a">Saltini, Claudio J. P.</subfield>
100	1	\$a Saltini, Claudio J. P., \$d 1935-	<subfield code="d">1935 - .</subfield>
245	1 0	\$a Afetividade e inteligencia /	</datafield>
260		\$a Rio de Janeiro : \$b DP&A, \$c 1997.	<datafield tag="650" ind1="0" ind2="4" >
300		\$a nv. : \$b il.	<subfield code="a">Psicologia</subfield>
504		\$a Inclui bibliografia.	<subfield code="x">Afeto.</subfield>
505		\$a Conteudo: v.1. A Emocao na Educacao.	<subfield code="x">Emoções.</subfield>
650	0 4 0 4 0 4	\$a Afeto \$b Psicologia \$a Emoções \$a Psicologia escolar	<subfield code="x">Psicologia escolar.</subfield>
697		\$a Psicopedagogia \$a Psicologia educacional \$a Aprendizagem	</datafield>

Fonte: Elaborado pelo autor.

A *tag datafield* aplicada na conversão da figura 16 possui as propriedades 100 e 650 que, por sua vez, encapsulam os subcampos \$a e \$d aplicados ao Nome Pessoal e os subcampos \$a e \$x aplicados ao subcampo 650. Os vínculos entre os dois campos representados foram minimamente codificados em consonância aos elementos básicos de descrição presentes no formato MARC à esquerda. Apesar de a conversão realizada ter privilegiado somente estes dois campos, poder-se-ia ampliar a relação de subcampos caso se faça necessário.

Esta é uma importante característica do formato MARC, qual seja, a extensão dos elementos aplicáveis à descrição. Atributo perceptível se comparado a outros esquemas e vocabulários com finalidade semelhante.

Formatos com reduzida extensão descritiva como o DC, ocasionam perda de informações quando convertidas de volta ao esquema MARCXML. No entanto, segundo (ASSUMPCÃO, 2013; FLAMINO, 2006; GONZALES, 2014) e a exemplo da seleção

tratada no parágrafo anterior, é possível estabelecer ligações entre os pontos de acesso principais retornados pelo DC a esquemas como MADS e MODS.

A figura 17 estende, por meio de remissivas e outros metadados associados, a descrição de uma autoridade de Nome Pessoal, o que torna o MARCXML compatível com a extensão do MARC21 na cobertura dos pontos de acesso e adequado para o estabelecimento de equivalências semânticas com vocabulários reduzidos, a exemplo do DC.

Figura 17: Aplicação do MARCXML na codificação de dados de autoridade de Nome Pessoal e remissivas correspondentes.

100 1_ la Eco, Umberto ld 1932-2016	<code><marc:datafield tag="100" ind1="0" ind2="1"> <marc:subfield code="a">Eco, Umberto</marc:subfield> <marc:subfield code="d">1932-2016</marc:subfield> </marc:datafield></code>
400 1_ la Eko, Umberto ld 1932-2016	<code><marc:datafield tag="400" ind1="1" ind2="0"> <marc:subfield code="a">Eko, Umberto</marc:subfield> <marc:subfield code="d">1932-2016</marc:subfield> </marc:datafield></code>
400 1_ la Eko, Humberto ld 1932-2016	<code><marc:datafield tag="400" ind1="1" ind2="0"> <marc:subfield code="a">Eko, Humberto</marc:subfield> <marc:subfield code="d">1932-2016</marc:subfield> </marc:datafield></code>
670 __ la Autor de: Le nom de la rose. 1983	<code><marc:datafield tag="670" ind1=" " ind2="0"> <marc:subfield code="a">Autor de: Le nom de la rose. 1983</marc:subf </marc:datafield></code>
670 __ la http://authorities.loc.gov 9/8/10 lb (d.n., rem.s.)	<code><marc:datafield tag="670" ind1=" " ind2="0"> <marc:subfield code="a">http://authorities.loc.gov9/8/10.</marc:subfi <marc:subfield code="b">(d.n.,rem.s.)</marc:subfield> </marc:datafield></code>

Fonte: elaborado pelo autor.

Os campos 400 na figura 17 representam remissivas para a entrada principal, ou ainda, formas variantes nas quais um nome pessoal também é conhecido. Suas inserções no catálogo de autoridades traz alguns aspectos imprescindíveis na recuperação dos registros, tais como: associação de dados de autoridade a registros bibliográficos, manutenção e consistência da base de dados.

Sobre o último aspecto, outros itens podem ser acrescentados, como o tratamento de variações de um mesmo nome, remoção das *stop-words* e dos sinais diacríticos. As abordagens semântica e sintática propiciada pela linguagem XML mostra-se capaz de alcançar estes objetivos em paralelo a algoritmos codificados para promover a

redução das inconsistências notacionais dos valores atribuídos aos campos e subcampos do formato MARC.

Dados adicionais destinados a atribuir papéis ao Nome Pessoal que figura como cabeçalho de entrada de um registro bibliográfico também são suportados pelo MARCXML. A exemplo de uma Entrada Secundária representada pelo campo 700 e subcampo \$e, que agrega atributos de qualificação como autor, tradutor, colaborador, organizador dentre outros. Tais atributos, ou instâncias da entidade representada, encontram nos esquemas MADS *namespaces* apropriados para descrever e representar no catálogo de autoridades, tópicos adicionais relevantes que caracterizam as autoridades.

6.4 Codificação em MADS

Conforme assinalado no capítulo anterior, o esquema MADS amplia os metadados destinados a representar dados de autoridade, o que inclui outras categorias de entidades além dos nomes pessoais conforme detalhado no capítulo 4.3 da página 42.

As propriedades constantes nos *namespaces* comportam atributos que visam relacionar entidades e manifestações na perspectiva do paradigma entidade-relacionamento segundo terminologia adotada pelos FRBR.

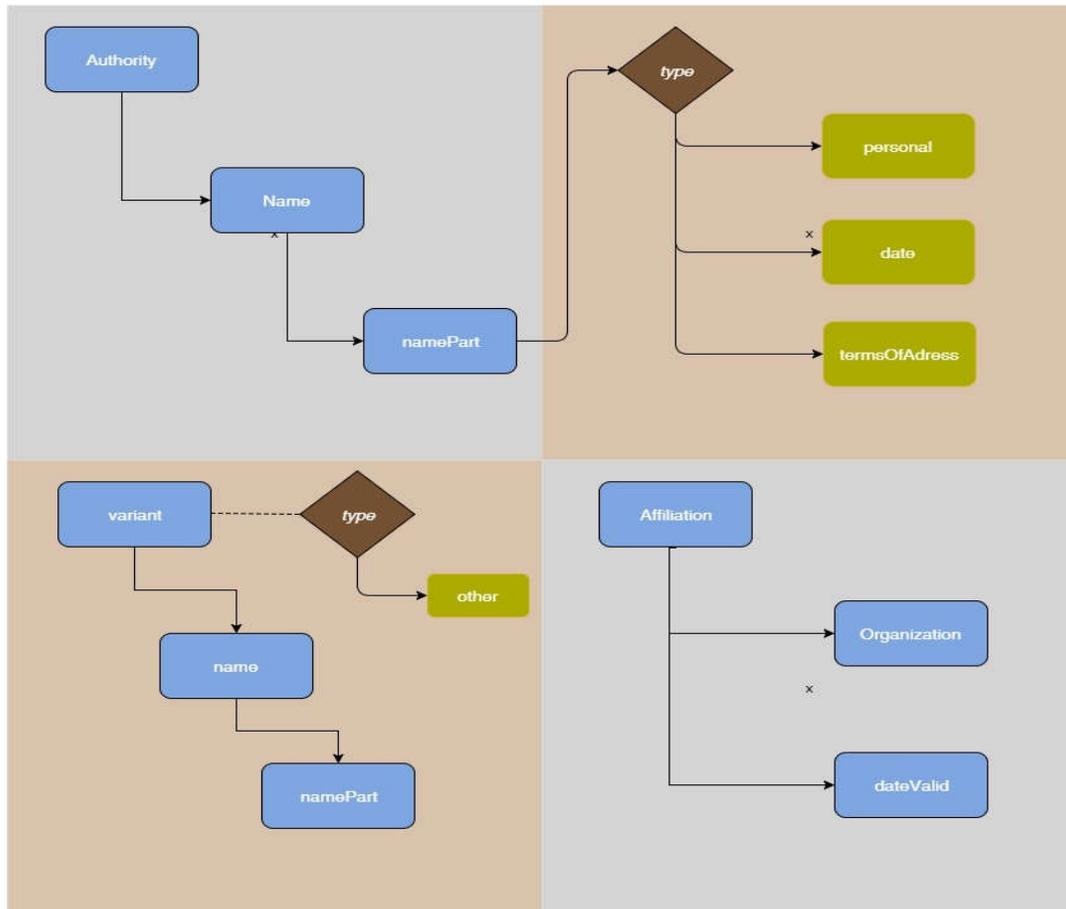
Para a codificação em MADS foi considerada a hierarquia das *tags* que integram o esquema com vistas a melhor representar os elos que associam as instâncias às propriedades da entidade representada.

Os vínculos internos ao esquema MADS são encontrados no MODS e MARCXML, e na figura 18 estão representados pela propriedade *type*, uma importante instância que atua como elo de ligação entre o nome pessoal e seus atributos correlatos.

No crosswalking realizado, foi verificado que a exploração destes atributos, como: *personal*, *date* e *termOfAddress*, constituem importante fator de preenchimento de lacunas aqui entendidas como eventuais perdas de informação da passagem de MARC para os esquemas XML.

Subtipos associados a *Affiliation* incrementam, em especificidade, o rigor da representação deste nome pessoal. No esquema da figura 18, foram arrolados, respectivamente *organization* e *dateValid*:

Figura 18: Esquema MADS baseado em autoridade de Nome Pessoal.



Fonte: Elaborado pelo autor.

O elemento raiz de um esquema MADS, conforme assinalam Assumpção e Santos (2013, p. 117), possui dois valores definidos a partir da quantidade de registros representados. O elemento denominado *<mads>* é utilizado para encapsular *tags* que descrevem um único registro, já o elementos *<madsCollection>* é utilizado como raiz quando se pretende descrever mais de um registro. O modelo esquemático da figura 18 estabelece como elemento de entrada da descrição a *tagAuthority* (MADS, 2016), cuja inserção na estrutura sintática do esquema MADS é assinalada por Guenther (2007) como uma propriedade que permite associar a entidade aqui representada pelo nome pessoal às suas formas variantes.

A codificação dos metadados representados na hierarquia da figura 18 em MADS ratifica este aspecto. Seus atributos relacionais agregam ao nome pessoal tipos variantes de dados, o que acrescenta e amplia sua semântica e legibilidade, aspectos considerados intrínsecos aos vocabulários e esquemas de metadados escritos em XML.

Figura 19: Aplicação do MADS na codificação de dados de autoridade de Nome Pessoal.

100 1 \$a Chauí, Marilena de Sousa, \$d 1941-	<mads> <authority> <name> <namePart> Chauí, Marilena de Sousa </namePart> <namePart type="date">1941-</namePart> </name> </authority>
400 1 \$a De Souza Chauí, Marilena, \$d 1941-	<variant type="other"> <name> <namePart>De Souza, marilena</namePart> </name> </variant>
400 1 \$a Souza Chauí, Marilena de, \$d 1941-	<variant type="other"> <name> <namePart>Souza Chauí, Marilena de</namePart> </name> </variant>
400 1 \$a Chauí, M. S. \$q (Marilena de Souza), \$d 1941-	<note type="history">Filósofa e escritora brasileira</note> <affiliation> <organization>Universidade de São Paulo</organization> <dateValid>19-?</dateValid> </affiliation>
400 1 \$a Chauí, Marilena de Sousa, \$d 1941-	</mads>
400 1 \$a Sousa Chauí, Marilena de, \$d 1941-	
670 \$a CatCD, LC Aut., (May 1999)	

Fonte: Elaborado pelo autor.

Verifica-se na codificação da autoridade pessoal na figura 19, que a cada nível hierárquico o esquema MADS agrega atributos relacionados ao subelemento de nível imediatamente superior. Esta característica traz maior legibilidade aos dados que se relacionam à autoridade representada. Em *affiliation*, por exemplo, todos os subcampos MARC que se destinam a identificar a filiação institucional da autoridade podem aqui ser arrolados, conforme será tratado mais adiante.

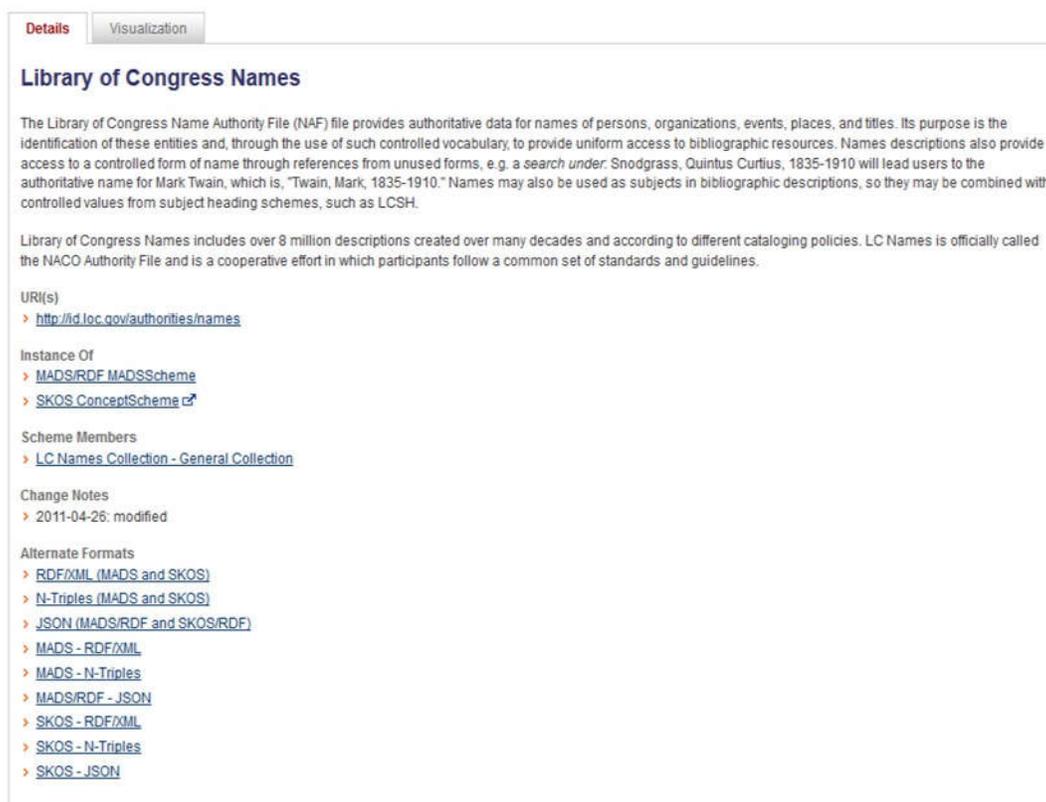
O caráter extensível e a flexibilidade do esquema MADS, por outro lado, proporcionam aos gestores de conteúdos, às bibliotecas e centros de documentação, roteiros customizados que permitem selecionar quais metadados são úteis para representar um catálogo de autoridades. Sobre esta característica, El Sherbini e Klim (2004, p. 241) lembram que padrões de metadados podem ser de terceiros e, nesta condição, mantidos em bases de dados e/ou repositórios ou podem, ainda, serem objetos de desenvolvimento da própria instituição que os cria e os mantém como vocabulários próprios.

A representação de um mesmo cadastro de autoridade de nome pessoal em formatos diferentes de metadados, amplia consideravelmente a interoperabilidade

semântica e o acesso descentralizado aos cabeçalhos se nomes, o que confere suporte às diversas plataformas tecnológicas destinadas à catalogação.

A definição da arquitetura de informação com ênfase nos esquemas de metadados baseados em XML são fundamentais para provimento de padrões flexíveis e adaptáveis a mudanças ambientais. Essas mudanças envolvem desde alterações nas políticas de informação das instituições a versionamento de *softwares* de automação, o que diminui consideravelmente a possibilidade de perdas de acesso a registros uma vez que sistemas de automação atuais possuem arquiteturas totalmente baseadas na *web* e dotadas de suporte para a linguagem XML.

Figura 20: Interface de busca no catálogo de autoridade da *Library of Congress*



Fonte: MADS (2016)

A figura 20 apresenta a interface de busca do catálogo de autoridades de nomes da Biblioteca do Congresso Americano. Ela ilustra a tendência contemporânea de fazer convergir seus formatos disponíveis para interoperabilidade, o que atende a demandas de compartilhamento de recursos para diversas plataformas voltadas à automação de bibliotecas e centros de documentação.

Em consonância ao esquema MADS representado na figura 16 da página 68 e o caráter extensível de sua descrição anteriormente tratado, os catálogos de acesso público ou *online access catalogue* (OPAC) tem agregado informações adicionais sobre a autoridade representada, tais como biografias, filiação, datas relacionadas dentre outras.

A conversão não se limita apenas a transportar campos e subcampos para esquemas de metadados da *web*, mas possibilita abordar o tratamento semântico do formato MARC em contexto multiplataforma, o que envolve desde aplicações *desktop* a dispositivos móveis.

Na conversão do campo 100 do formato MARC para o esquema MADS na figura 19, a abordagem semântica está presente não apenas na terminologia em XML, mas na composição hierárquica das propriedades associadas ao nome. A *tag affiliation*, por exemplo, atua como *container* para atributos da autoridade de nome, o que inclui subelementos do campo 373 – Grupo Associado. Este campo do formato MARC reflete a amplitude e complexidade dos campos variáveis que estão, em maior ou menor número, presentes nas bases de autoridades de nomes.

A conversão representada na figura 21 abstrai a complexidade destes campos por meio de *namespaces* apropriados, o que inclui também formatos válidos para exibição de datas por meio da *tag datteValid*.

Figura 21: Codificação de subelemento da autoridade de nome.

<affiliation>	
373	<affiliation>
373\$a	<position>
373\$s	<dateValid encoding="iso8601" point="start">
373\$t	<dateValid encoding="iso8601" point="end">
371	<affiliation>
	<address>
371\$a	<street>
371\$b	<city>
371\$c	<state>
371\$d	<country>
371\$e	<postcode>
371\$m	<email>

Fonte: MADS(2016)

Os subtipos associados a *address*, na figura 21, arrolam os metadados necessários para registro de endereço da autoridade, cujo equivalente em MARC é representado pelo campo 371 – *Address*. A diversidade de campos disponíveis em MARC e aptos a serem selecionados para integrar um catálogo de autoridade auxilia sobremaneira na definição de uma arquitetura de cunho relacional e distribuída, na medida em que se procede a codificação em esquema XML adaptado a representar na extensão e alcance necessários os atributos, as propriedades e as relações das autoridades de registros bibliográficos.

Em paralelo a campos e subcampos para descrição de dados adicionais conforme demonstrado anteriormente, outras *metatags* em MADS em destaque para uso opcional operacionalizam o registro em lide, por meio de relações estabelecidas entre si direta ou indiretamente. O quadro 10, agrega os campos destinados ao estabelecimento de vínculos à entidade representada, com *namespaces* que ocupam na estrutura em MADS a mesma posição hierárquica.

Quadro10: Descrição dos *namespaces* em MADS e respectivas equivalências MARC.

Namespace MADS	Descrição	Equivalente em MARC
<affiliation>	Permite codificar afiliação e informações de contato para uma pessoa ou organização.	373
<classification>	Inclui uma classificação ou número de chamada a um recurso.	053, 055, 060, 065, 070, 080, 082, 083, 086, 087
<fieldOfActivity>	Inclui um campo de atuação ou negócio, área de expertise, ou área de competência de uma pessoa ou organização.	372
<identifier>	Inclui um identificador associado com uma entidade descrita no registro.	010, 016, 020, 022, 024, 035
<language>	Inclui o idioma principal do registro; possui como subelementos as tags <languageTerm> e <scriptTerm>.	377
<note>	Inclui informação geral (equivalente a Nota Geral) se nenhum atributo do tipo <i>type</i> for fornecido.	680
<url>	Inclui endereços da internet atribuídos ao recurso.	856
<extension>	Utilizado para cadastrar elementos locais do criador dos dados.	9XX
<recordInfo>	Elemento <i>container</i> cujos subelementos relacionam informações necessárias ao processamento dos dados.	<i>Leader</i>

Fonte: Elaborado pelo autor.

As descrições dos elementos presentes no quadro 10, representam subcampos em MARC aptos a serem utilizados não apenas na catalogação de autoridades de nomes pessoais, mas também no estabelecimento de vínculos com demais registros associados o que também inclui assuntos. É verificado que, apesar da codificação da autoridade de nome pessoal na figura 21 utilizar os *namespaces* do quadro 10, o uso do esquema MADS comporta também registros bibliográficos, seus assuntos e os vínculos entre ambos e demais entidades do domínio a representar.

Na conversão dos registros de autoridade foram identificados diversos vínculos, ou *linked data*, destinados a estabelecer elos de ligação entre autoridades e recursos bibliográficos. Alguns *namespaces* como `<recordContentSource>` e `<recordIdentifier>`, possuem tarefas bem definidas: o primeiro especifica o código da instituição que registrou o termo autorizado, campo 040 do MARC. Já o segundo, operacionaliza o intercâmbio através de um código identificador próprio, campo 001 do MARC.

O formato MARC possui o campo 856, representado pelo MADS através da *tagurl* no quadro 10. Propriedades adicionais a links relacionados a um registro ou mesmo a uma autoridade podem ser agregados a esta *tag*. Um exemplo é a lista de indicadores que apontam para a localização do recurso descrito, que inclui *Email*, *FTP*, *login* remoto, *dial-up*, *http* e outros métodos de acesso. O campo 856 perfaz um total de 8 indicadores aptos a serem codificados como propriedades da *tag url* em MADS.

Figura 22: Aplicação do MADS na codificação dos atributos relacionados à nome pessoal.

<p>100 \$a Pinto, Virginia Bentes</p> <p>400 \$a Bentes Pinto, Virginia</p> <p>670 \$a Organizadora de: Aplicabilidades metodológicas em Ciência da Informação, 2015</p> <p>670 \$a BN on-line (18/05/2009)</p>	<pre> <mads> <authority> <name type="personal" authority="naf" > <namePart>Pinto, Virginia Bentes</namePart> </name> </authority> <related type="equivalent"> <name type="personal"> <namePart> Bentes Pinto, Virginia</namePart> </name> <role> <roleTerm type="text">Organizadora</role> <occupation>Pesquisadora</occupation> </role> </related> <related type="broader"> <source>Organizadora de: Aplicabilidades metodológicas em Ciência da Informação </source> <geographic>Fortaleza, CE</geographic> </related> </authority> </mads> </pre>
--	--

Fonte: Elaborado pelo autor.

Os metadados adicionais previstos no esquema MADS que agregam àqueles atributos destinados a representar as formas variantes de nomes, já representado na figura 19 da página 72, são seguidos de outros que tem como objetivos o estabelecimento de vínculos entre as entidades que perfazem o domínio bibliográfico de um acervo. Estes vínculos permitem ao usuário do catálogo navegar pelas obras cujo nome da autoridade se encontra atrelado a diversos papéis *role*, a exemplo da figura 22.

A tag *roleTerm* agrega duas propriedades mandatórias: *code* e *text*. Ambas especificam o tipo de dados aceitos como valor, sejam eles codificados e textuais respectivamente. A referida especificação foi tratada no capítulo 3.2.1 da página 34, ocasião em que se abordou o DTD e seu papel validador. No entanto, como se observa nos esquemas MADS e MODS, a validação pode ser aplicada no corpo do próprio documento XML com vistas a minimizar perdas de dados úteis ao catálogo. Um exemplo baseado na figura 22 ilustra esta situação: abaixo da tag `<roleTerm="text">Organizadora</role>` poder-se-ia adotar outra do tipo `<roleTerm="code">Org.</role>`. Esta última representa o subcampo \$e ou \$4 da Entrada Secundária de Nome Pessoal, ou campo 700 no formato MARC. Trata-se dos termos de relação atribuídos à autoridade, em que a propriedade *code* possui acrônimos padronizados e extraídos das bases de autoridades.

Formas variantes de Nome Pessoal como Entrada Secundária tem ainda a propriedade *equivalent*, que na posição hierárquica do MADS comporta quantos pontos de acesso autorizados sejam necessários para representar as formas pelas quais um nome é conhecido e citado nas mais diversas manifestações. Cabe reiterar que a entrada principal de uma autoridade é sua forma mais conhecida.

Alguns dos campos representados no quadro 11 são classificados como repetitivos (R) ou não repetitivos (NR), ou seja, podem ocorrer uma única vez ou mais de uma vez no formato MARC ou no esquema MADS. No entanto, sua inclusão na descrição dos recursos está vinculada ao conteúdo, cuja presença pode ser encontrada nas diversas manifestações, sejam elas geográficas, de gênero, de ocupação dentre outras.

A ocorrência simultânea dos mesmos *namespaces* tanto em MADS quanto em MODS foi verificada na medida em que se procedeu as conversões, uma vez que tanto em autoridades de nomes quanto em autoridades de assuntos o uso de subelementos como o geográfico e temporal se inserem em contextos variáveis de representação.

6.5 Codificação em MODS

O esquema MODS pode ser aplicado para descrever diversas facetas de um recurso bibliográfico a partir dos campos e subcampos MARC correspondentes. Foi observado, nas conversões realizadas, que a seção do esquema destinada à catalogação de assuntos foi capaz de estender o alcance dos termos tópicos conforme será visto neste capítulo.

O quadro 11 é resultado do mapeamento e conversão de alguns dados de autoridade de assunto, representados pela sintaxe MODS na coluna da direita:

Quadro11: Conversão de campos de assunto MARC para MODS.

Subject	<subject>
600, 610, 611, 630, 648, 650, 651	<subject authority> <i>with</i> attribute:
650 \$a \$b \$c \$d	<topic>
653 \$a ind2 = 0	<name type="personal"> <namePart>
653 \$a ind2 = 2	<name type="corporate"> <namePart>

Fonte: MARC TO MODS (2014)

Para cada campo do formato MARC, o esquema MODS possui seu equivalente com os respectivos indicadores e subcampos, codificados em sintaxes legíveis por humanos e máquinas.

A hierarquia inerente ao esquema XML tem como característica prover flexibilidade a hierarquia implícita no formato MARC. Para os assuntos tópicos representados pelos campos que seguem do 600 ao 610, por exemplo, verifica-se o correspondente *subject authority*, cujo desdobramento acomoda atributos relacionados ao elemento de nível superior *subject*.

Propriedades como *personal* e *corporate* não correspondem a subdivisões hierárquicas, mas a instâncias da *tag name*. O mesmo ocorre com a *tag subject* que contempla propriedades que atuam como instâncias do termo tópico.

Os indicadores, por sua vez, podem sinalizar a presença de subelementos em cadeia hierárquica. Seu uso no formato MARC está relacionado a regras a serem seguidas pela entrada em lide. Para o campo 650 por exemplo, o indicador 1 sinaliza que o assunto tópico é primário, ou termo mais relevante. O indicador 2, por sua vez, sinaliza o caráter secundário de um assunto ou termo tópico, o que não prescinde de sua relevância, mas não o situa como elemento de topo na hierarquia de termos.

A codificação em MODS de uma relação de termos tópicos não ocorre em prejuízo a hierarquia presente em MARC, pelo contrário, o uso das propriedades inerentes a este esquema descritivo não apenas consolida a hierarquia como a qualifica por meio dos atributos adicionais.

Assim como o MADS, o esquema MODS pode ser aplicado para descrever recursos bibliográficos no todo, a exemplo da figura 23 na página 83. No entanto, é nos subelementos do *namespace subject* que se encontram atributos que ampliam consideravelmente as opções de consulta de um assunto tópico. A mesma relação ampliada de atributos descritivos também se faz presente em MADS ao se estabelecer os pontos de acesso para nomes pessoais demonstrado no capítulo anterior.

Os *namespaces* em MODS e suas respectivas equivalências em MARC foram mapeados no quadro 12, cujas descrições apontam para a função destes pontos de acesso na descrição como termo tópico, secundárias e remissivas. Observa-se, pois, que as facetas de uma dada autoridade em MARC está, na sua literalidade, presente em MODS.

Quadro 12: Descrição dos *namespaces* em MODS e respectivas equivalências em MARC.

Namespace MODS	Descrição	Equivalente em MARC
<topic>	Termo genérico que descreve o conteúdo a ser representado.	650 e 150
<geographic>	Designação de lugar como assunto ou subdivisão de termo tópico.	651
<titleInfo>	Título usado como assunto.	630
<name>	Um nome usado como assunto.	600
<genre>	Gênero ou parte de um termo tópico.	008
<hierarchicalGeographic>	Nome Geográfico em forma hierárquica.	752, 662
<cartographics>	Dados de mapas ou cartas indicativos de cobertura espacial.	034
<geographicCode>	Código de área geográfica associada a um recurso.	043
<occupation>	Termo referente a ocupação referida nos conteúdos dos assuntos descritos.	<i>Leader</i>

Fonte: Elaborado pelo autor.

Ao proceder a conversão de uma autoridade de assuntos com as respectivas variáveis geográficas, temporais e de gênero para vocabulários de cunho genérico como o DC, fica evidente a lacuna semântica deste último caracterizada pela ausência dos metadados necessários a uma exaustiva representação das autoridades concernentes ao domínio bibliográfico. A abordagem genérica do DC presta-se à catalogação de recursos de informação na *web* que não comportam em especificidade a demanda de metadados com propriedades, atributos e relações de um domínio de assunto, ou as particularidades de uma autoridade de nome.

O quadro 13, ilustra a equivalência de alguns elementos DC com metadados MODS aplicados à descrição de assuntos. A finalidade desta conversão, assim como a escolha do DC como exemplo, tem como premissa fundamental estabelecer uma amos-

tra comparativa da exaustividade de pontos de acesso e especificidade na cobertura de assuntos conforme discutido anteriormente.

Quadro 13: Subelementos MODS expressos em *Dublin Core*

MODS <subject> subelements	DC elements
<topic> <name> <titleInfo> <occupation>	<dc:subject>
<geographic> <temporal> <hierarchicalGeographic>	<dc:coverage>
<genre usage="primary"> <genre type="style" displayLabel="culture"> <genre type="style" displayLabel="style">	<dc:type>
<cartographics> <geographicCode>	[no field suggested]

Fonte: LIBRARY OF CONGRESS(2018)

Embora os *namespaces* MODS do quadro 13 sejam subelementos de *subject*, foi verificado nesta conversão da LC que sua aplicabilidade as situam em categorias superior, ou *top level elements*, em que a subdivisão hierárquica se consolida por meio dos atributos relativos ao assunto que se deseja representar. Esta característica é verificada em ambos os esquemas e ilustrado na figura 16 da página 67.

Em relação ao MODS, o elemento *genre* no quadro 13 comporta atributos como: *type*, *authority*, *authorityURI*, *valueURI*, *usage*, *displayLabel*. (LIBRARY OF CONGRESS, 2018). Estes, por sua vez, se associam a propriedades relativas ao elemento *subject* ou termo tópico representado. À *displayLabel* as propriedades *style* e *culture*, por exemplo, acrescentam aspectos específicos ao termo tópico, ao mesmo tempo em que *usage* vincula-se a propriedade *primary* em referência ao assunto principal tratado em *subject*.

Estes são apenas alguns dos atributos e propriedades passíveis de serem incorporados ao esquema MODS. Contudo, ao consultar um registro bibliográfico com-

pleto em OPAC e visualizar o resultado gerado, é possível identificar que outras áreas de descrição seguem os mesmos paradigmas de expressividade semântica dos metadados. Verifica-se que a paridade das áreas de descrição das entidades do domínio bibliográfico em MARC é igualmente reproduzida em MODS com perda mínima de informação e dotada de legibilidade tanto por humanos quanto por máquinas.

Figura 23: Esquema MODS de registro consultado no catálogo da LOC

```

▼<titleInfo>
  <title>Preparation for teaching</title>
  <subTitle>a standard course for teacher training</subTitle>
</titleInfo>
▼<name type="personal" usage="primary">
  <namePart>Oliver, Charles Alexander. [from old catalog]</namePart>
</name>
<typeOfResource>text</typeOfResource>
▼<originInfo>
  ▼<place>
    <placeTerm type="code" authority="marccountry">pau</placeTerm>
  </place>
  ▼<place>
    <placeTerm type="text">Philadelphia</placeTerm>
  </place>
  <publisher>The Westminster press</publisher>
  <dateIssued>1909</dateIssued>
  <issuance>monographic</issuance>
</originInfo>
▼<language>
  <languageTerm type="code" authority="iso639-2b">eng</languageTerm>
</language>
▼<physicalDescription>
  <form authority="marcform">print</form>
  ▼<extent>
    x p., 1 l., 132 p., 1 l. illus. (maps, plan) 19 cm.
  </extent>
</physicalDescription>
▼<note type="additional physical form">
  Also available in digital form on the Internet Archive Web site.
</note>
▼<subject authority="lcsht">
  <topic>Sunday schools</topic>
</subject>
<classification authority="lcc">BV1533 .06</classification>
▼<location>
  <url displayLabel="electronic resource" usage="primary display">http://hdl.loc.gov/loc.gdc/scd0001.0014745224A</url>
</location>
<identifier type="lccn">09014939</identifier>
<identifier type="hdl">hdl:loc.gdc/scd0001.0014745224A</identifier>

```

Fonte: LIBRARY OF CONGRESS (2018)

No esquema MODS da figura 23, os dados de publicação, descrição física e de autoridade de nome pessoal são acrescidos a *namespaces* destinados ao estabelecimento de vínculos internos na coleção de uma biblioteca ou, ainda, vincular dados com outras instituições. A individualização do registro em uma coleção e sua vinculação à unidade de informação são efetivadas por meio da *tag identifier*, cuja propriedade *lccn* e *hdl* agregam códigos identificadores de recursos com atribuições similares aos números de tombo no interior das coleções.

Os metadados que ocupam posição superior na hierarquia do esquema MODS não prescindem de regras formais orientadas a uma catalogação mais exaustiva, ou de terceiro nível. Instruções claras sobre a adoção e exclusão de um ou outro *namespace* em bases de autoridades envolvem aspectos atinentes ao contexto de uso do acervo, dimensões do mesmo, assuntos tratados e nível de conhecimento do usuário somente para citar algumas variáveis.

Sobre este aspecto e relativamente à vinculação dos assuntos aos suportes físicos de um acervo, a codificação dos subelementos listados no quadro14 orientam, no ato da conversão, a escolha de alguns metadados aptos a gerar pontos de acesso na interface com o usuário.

Quadro14: Subelementos MODS aplicados à descrição de assunto.

<i>Namespace</i>	Subelementos	Definição	Aplicado a
<titleInfo>	<title>	Palavra, frase, caractere ou grupo de caracteres atribuído a um recurso e destinado a nomeá-lo.	Livros, coletâneas, eventos, coleções, periódicos dentre outros recursos.
	<subTitle>		
	<partNumber>		
	<partName>		
	<nonSort>		
<hierarchicalGeographic>	<continent>	Um nome geográfico dado em forma hierárquica a um recurso.	Locais de publicação, realização de eventos e informações relativas a mais de um local de edição em obras editadas mais de uma vez.
	<country>		
	<region>		
	<state>		
	<territory>		
	<county>		
	<city>		
	<citySection>		
<cartographics>	<coordinates>	Mapas ou cartas cartográficas indicativas de cobertura espacial.	Atlas geográficos, mapas de relevo e demais recursos que se utilizem de convenções cartográficas.
	<scale>		
	<projection>		

Fonte: Adaptado de LIBRARY OF CONGRESS(2017)

Conforme assinalado anteriormente, alguns subelementos estão passíveis de serem selecionados para compor pontos de acesso orientados ora ao assunto tratado, ora ao suporte que o contém. Em coleções de periódicos alguns metadados são de aplicabi-

lidade obrigatória, como *title* e *partNumber*. O primeiro atinente ao título, e o segundo indicativo de seção no todo ou em parte, como um fascículo de uma coleção ou o número de um artigo de periódico.

Para as coleções de periódicos ou cabeçalhos de assuntos que lidam com dados geográficos como assunto principal em que suas disposições no catálogo se fazem imprescindíveis ao usuário, os subelementos de *hierarchicalGeographic* presentes no quadro 14 ilustram a extensão descritiva do MODS para esta faceta.

Em observância à catalogação de assunto tópico, a figura 24 resultou da conversão do campo 150 e suas remissivas em um esquema MODS adaptados para interoperar com outras manifestações e entidades do domínio representado por meio da inserção do atributo *valueURI* discutido a seguir.

Importa destacar o uso da propriedade *public general note* do atributo *note type* na figura 24, cujo campo análogo em MARC é o 500 destinado a notas gerais, que aqui tem a premissa de agregar à lista de Termos Tópicos uma definição sintética extraída de uma base de autoridades.

Relativamente ao uso de notas, o formato MARC agrega diversas facetas ao assunto representado em campos como 504 – Nota de Bibliografia, 505 – Nota de Conteúdo, 507 – Nota de Escala para Material Gráfico, 520 – Nota de Resumo, dentre outras. Os equivalentes em MODS, na documentação da (LIBRARY OF CONGRESS, 2018), preservam as dimensões genérica e específica dos conteúdos das notas sem perdas de informação.

A propriedade *statement of responsibility* ou indicação de responsabilidade, por exemplo, é mais compatível com a *tag titleInfo* quando o esquema MODS é usado para o estabelecimento de ligações entre autor e obra, ou pessoa e entidade em alusão à terminologia dos FRBR.

Notas históricas podem ser incorporadas a *note type* por meio de uma propriedade denominada *historical*, cuja inserção no catálogo, em paralelo a *public general note*, torna mais representativo o assunto em tela em função da especificidade que propicia o usuário navegar de notas gerais a específicas.

Figura 24: Aplicação do MODS na codificação de identificadores como remissivas.

<p>150 \$a Comunismo \$z (subd. geog.)</p> <p>450 \$a Bolchevismo</p> <p>450 \$a Marxismo</p> <p>450 \$a Bolshevism</p> <p>450 \$a Communist movements</p> <p>450 \$a Leninism</p> <p>450 \$a Maoism</p> <p>450 \$a Marxism</p> <p>550 \$w g \$a Coletivismo</p> <p>550 \$w g \$a Totalitarismo</p> <p>550 \$w h \$a Economia marxista</p> <p>550 \$w h \$a Movimentos anticomunistas</p> <p>550 \$w h \$a Propaganda comunista</p> <p>550 \$w h \$a Socialismo e juventude</p> <p>651 \$a Brasil</p> <p>670 \$a LC (2001)</p> <p>680 \$i Aqui entram trabalhos que tratam do movimento revolucionário ou ideologia inspirada em Marx e que defende a abolição da propriedade privada, ditadura do proletariado e a extinção gradual do Estado. Nos dias de hoje o movimento comunista é caracterizado pela propriedade coletiva dos recursos de produção.</p> <p>750 0 \$a Communism</p>	<pre> <mods> <subject authority="bn"> <topic> Comunismo</topic> <topic> <topic valueURI="">Coletivismo</topic> <topic valueURI=""> Totalitarismo </topic> <topic valueURI=""> Economia marxista </topic> <topic valueURI="">Movimentos anticomunistas</topic> <topic valueURI="">Propaganda comunista </topic> <topic valueURI="">Socialismo e juventude </topic> <topic valueURI="">Marxism</topic> </topic> <note type="public general note">Aqui entram trabalhos que tratam do movimento revolucionário ou ideologia inspirada em Marx e que defende a abolição da propriedade privada, ditadura do proletariado e a extinção gradual do Estado. Nos dias de hoje o movimento comunista é caracterizado pela propriedade coletiva dos recursos de produção. </note> <geographic>Brasil</geographic> <temporal>Século XX</temporal> </subject> </mods> </pre>
--	--

Fonte: Elaborado pelo autor.

Um identificador único de recurso ou URI pode ser tratado como elemento chave para nomear e vincular unidades semânticas do domínio de assunto no ambiente distribuído da *web*. A LC define o atributo *valueURI* como elemento destinado a identificar o termo ou valor controlado designado pela autoridade responsável pela manutenção do vocabulário (LIBRARY OF CONGRESS, 2017)

Outro identificador com função semelhante, mas com aplicabilidade orientada a termos tópicos presentes em bases de autoridades é o atributo *authorityURI*. Este pode ser aplicado simultaneamente às *tags subject authority* e em conjunto com *valueURI*.

A conversão da figura 25, destitui do elemento raiz *subject* do esquema MODS quaisquer atributos adicionais como os mencionados no parágrafo anterior. Esta característica tem o condão de estabelecer duas estratégias de codificação de metadados orientados ao paradigma entidade relacionamento: a primeira quando existe a necessi-

dade de vincular termos tópicos, representados pela *tag subject*, às bases de autoridades que os contem.

A segunda, de aplicar os termos repetitivos a múltiplas bases de autoridade que não necessariamente utilizam as mesmas URI's para seus recursos cadastrados.

O consórcio *w3c* no relatório de seu grupo de trabalho relativo a URI, alerta para limitações de interoperabilidade baseada na inserção da arquitetura e sintaxe das URI's em comunidades abertas da internet. As plataformas de intercâmbio de informações que se utilizam deste identificador bem como do *Uniform Resource Locator* (URL) e do *Uniform Resource Name* (URN) atendem a finalidades de acesso e recuperação de recursos em comunidades especializadas que detém o domínio da terminologia representada nestas arquiteturas.

Contudo, na modelagem baseada no paradigma entidade relacionamento e adotada pelos FRBR, as lacunas semânticas apontadas acima são abstraídas por esquemas de metadados de cunho relacional.

A título das conversões apresentadas e discutidas anteriormente, a figura 25, agrega ao assunto tópico os subelementos em MARC que se fazem necessários à sua representação no catálogo, além de servir como pontos de acesso para busca e recuperação de recursos bibliográficos a partir da consulta à base de autoridades.

Figura 25: Conversão para subelementos de *subject* em MODS.

<p>611 24 \$a Congresso Brasileiro de Biblioteconomia e Documentação</p> <p>\$n (2. :</p> <p>\$d 1994 :</p> <p>\$c Belo Horizonte , MG)</p>	<pre> <subject> <topic>Gestão do conhecimento</topic> <topic>Recuperação da informação</topic> <topic>Bibliotecas públicas</topic> <topic>Repositórios institucionais</topic> <hierarchicalGeographic> Belo Horizonte, MG </hierarchicalGeographic > <temporal>1994</temporal> <name type="conference"> <namePart>Congresso Brasileiro de Biblioteconomia e Documentação </namePart> </name> <genre authority="loc">Evento científico</genre> <titleInfo> Anais do 2§ Congresso Latino-Americano de Biblioteconomia e Documentação e do 17§ Congresso Brasileiro de Biblioteconomia e Documentação. </titleInfo> </subject> </pre>
--	--

Metadados que expressam relações associativas entre entidades do domínio de assunto tratados em registros bibliográficos são abordados pelos FRSAD no contexto puramente relacional. O desdobramento da *tag* `<hierarquicalGeographic>` na figura 25, destina-se ao estabelecimento da relação associativa entre o local de realização do evento e o nome de evento tratado como assunto; campo 611 do MARC. Entretanto, embora situado como elemento filho de *subject*, quando existe a necessidade de tratar nomes geográficos como assunto de um recurso, *hierarquicalGeographic* encapsula metadados associativos como: *continent*, *country*, *region*, *state* dentre outros.

A despeito de a LC situar no esquema MODS o subelemento *hierarquicalGeographic* como “[...] forma hierárquica de nome de lugar” (LIBRARY OF CONGRESS, 2018), a sua aplicabilidade orientada ao conteúdo representado termina por deslocar o paradigma relacional presente no domínio bibliográfico para a estrutura hierárquica intrínseca a esquemas XML como o MODS.

Um exemplo pode ser visualizado, comparativamente, a partir da figura 24, em que foi utilizada a *tag* `<geographic>` na conversão, e na figura 25 em que se utiliza a *tag* *hierarquicalGeographic*. Em ambos, o nome de lugar geográfico assume facetas distintas relacionadas ao assunto tópico no primeiro caso e a nome de evento no segundo.

É possível agregar à *tag* *subject* atributos destinados a individualizar tipos de suporte, aspecto compatível com os esquemas MODS em que seus *namespaces* são aplicáveis à catalogação de “objetos” ou recursos de informação. Neste interim, a (LIBRARY OF CONGRESS, 2018) sinaliza a existenciado atributo *manuscript* marcado com o valor “yes” ou “no” em `<typeOfResource>`, cuja finalidade é indicar se o tipo de recurso de alto nível está em forma de manuscrito.

Esta flexibilidade notacional verificada em MODS, permite inferir que documentos digitalizados podem ser diferenciados no interior do catálogo daqueles documentos que nasceram em meio digital. As demais propriedades presente no esquema também facilitam a representação das relações entre autor intelectual e suas respectivas obras conforme tem sido verificado nas conversões.

7 ANÁLISE DOS DADOS E DISCUSSÃO DOS RESULTADOS

Como resultado das estratégias de conversões de metadados do formato MARC para os esquemas MARCXML, MADS e MODS, foi verificado o valor semântico do paradigma relacional preconizado pelos FRBR como elemento chave para alcance de expressividade, coerência interna e sofisticação das relações sintáticas e semânticas tão caras à efetividade dos catálogos. Neste íterim, ratifica-se a observação de Fusco (2010, p. 173), quando assevera que: “A Modelagem Conceitual pode auxiliar e melhorar o processo de desenvolvimento de ambientes informacionais digitais baseados no conceito de catalogação e a sua utilização nos FRBR corroborou essa ideia [...]”. Na mesma esteira de pensamento e em abordagem baseada em ambientes digitais, Bezerra (2013, p. 152) ressalta o *status* que a modelagem conceitual tem assumido ao nível de convenções normativas neste cenário ubíquo:

Notamos ainda, o uso, cada vez mais generalizado, de modelos conceituais na área da documentação a fim de agregar maior precisão semântica aos ambientes digitais e permitir o tratamento por computador de forma mais sofisticada. Além do já consagrado modelo CIDOC - CRM da área de museologia, destaca-se o desenvolvimento da norma ISO 25964 - 1:2011 –*Thesauri for Information Retrieval* e do vocabulário *Expression of Core FRBR Concepts in RDF* (<http://vocab.org/frbr/core.html>). Tais aplicações nos levam a refletir sobre os rumos da Catalogação diante do ambiente digital e as potencialidades do modelo FRBR neste ambiente. (BEZERRA, 2013, p. 152)

A necessidade de reflexão proposta pelo autor no que concerne aos rumos da catalogação, vai ao encontro dos resultados aqui verificados, em particular nos estudos dos atributos das entidades do domínio bibliográfico e suas relações no interior do catálogo.

Estas relações de ordem sintática e semântica perfizeram o mapeamento e a codificação das entradas de nome pessoal e assunto tópico como pontos de acesso selecionados para a conversão. O estudo das unidades semânticas aparece como eixo da abordagem de cunho relacional e, portanto, particularmente relevante para o contexto da modelagem conceitual, o que vem ao encontro da reflexão Zafallon (2012, p. 76):

A teoria das unidades semânticas volta-se ao estudo de um conjunto de itens lexicais de um registro bibliográfico como sendo semanticamente relacionados, cujos significados são mutuamente interdependentes, e que juntos fornecem a estrutura conceitual de um determinado registro bibliográfico.

Uma dimensão pragmática desta abordagem relacional e verificada na pesquisa, foi a preservação das remissivas atribuídas ao termo principal, associada a múltiplas possibilidades de relacionamentos hierárquicos e horizontais. A extensão do

formato MARC, tão realçada no curso das conversões realizadas, foi ratificada pela flexibilidade e legibilidade dos esquemas XML adotados e adaptados para descrição bibliográfica.

Fator de relevância estratégica dos esquemas de metadados baseados na *web* como o MARCXML, MADS e MODS, é a consistência alcançada quando da atribuição de identificadores vinculados aos dados. Para Zavalina, Shakeri e Kizhakkethil (2015, p. 6), a consistência semântica refere-se a extensão para a qual os mesmos valores ou elementos são usados para representar conceitos similares.

No caso de assuntos tópicos, foi verificado que a despeito de seu caráter repetitivo no catálogo, suas especificidades e exaustividades não prescindem do uso dos mesmos homógrafos e sinônimas. Desde que estes estejam devidamente atrelados a suas respectivas remissivas.

Comparativamente a outros formatos aplicados a descrição de recursos da *web*, o DC foi abordado em conjunto com os esquemas XML estudados por se tratar de um vocabulário amplamente citado e empregado em ambientes informacionais digitais.

Desta comparação, conduzida na etapa empírica, foi verificado o valor estratégico e agregado a um catálogo de autoridades codificado em XML, quando os pontos de acesso ali presentes são associados a atributos que, além de vincular autoridades, tornam a navegação pelo catálogo mais flexível e intuitiva.

No que concerne a eventuais perdas de informação como resultado da conversão, a despeito das menções na literatura a respeito, foi verificado que os equivalentes em MADS e MODS propiciam a manutenção da integridade dos valores presentes no formato MARC base. O cotejamento entre os *top level elements* e seus respectivos subelementos com os campos e subcampos MARC, não apenas mantiveram a integridade dos valores presentes nos campos MARC em estudo como ampliaram as possibilidades de enriquecimento do catálogo de autoridades.

As estratégias de conversão aqui exploradas se tornaram mais viáveis a partir de uma ferramenta selecionada para operacionalizar a codificação dos campos e subcampos MARC: O *MARC Editor*. Neste particular, o portal da Biblioteca do Congresso Americano possui links para provedores de aplicações destinadas à catalogação e conversão de metadados, com documentações concisas para sua instalação e uso.

O paradigma entidade-relacionamento adotado pelos FRAD e FRASAD perpassou toda a etapa de modelagem e codificação dos esquemas MARCXML, MADS

e MODS, por meio do estabelecimento de papéis atribuídos às autoridades bem como a identificação de suas propriedades e vínculos. A robustez da modelagem conceitual proposta pelos FRBR e aplicada nesta pesquisa às autoridades de nomes e assuntos, foi verificada desde a terminologia adotada, a exemplo das designações latinas *nomen* e *thema*, para melhor representar o universo do arcabouço informacional presente em catálogos de autoridades, a exaustiva exemplificação dos atributos passíveis de serem incluídos no modelo.

O esquema METS tratado no capítulo 4.1 da página 38, embora não tenha sido abordado nas conversões realizadas nesta pesquisa, deve ser destacado como esquema XML apto a ser usado para catalogação de objetos digitais. Seu uso pode ser abordado em pesquisas futuras sobre conversões de metadados de autoridade por meio da incorporação de fragmentos dos esquemas MADS e MODS em sua estrutura. Procedimento análogo de encapsulamento de metadados XML de esquemas diferentes pode ser visto em Coyle (2010) e Fachin (2011), com particular ênfase ao RDF e OWL.

Para catálogos de autoridade que contemplem informações sobre direitos autorais, foi verificado que o esquema MODS pode atuar em conjunto com o elemento *rights* do DC, seja para prover informações adicionais de nome pessoal, também presente em MODS, seja para arrolar subelementos que melhor representem o recurso catalogado.

Vocabulários controlados disponíveis para consulta em linha tem se tornado comum na *web*, e o suporte a linguagens extensíveis como XML atuam como nodo central nas camadas de conversão para formatos de exibição em interfaces com o usuário. Sobre este aspecto, as tecnologias XSLT e CSS possuem marcadores apropriados para devolver ao usuário interfaces gráficas amigáveis.

8 CONCLUSÕES

Toda investigação científica é holística, posto que tem um ciclo dinâmico e retroalimentável, razão pela qual se faz necessário retornar à questão de partida e os objetivos perseguidos ao longo de sua realização. Evidencia-se que não se tem a intenção de repetir os achados da pesquisa tratados na análise dos dados e discussão dos resultados, porém, salienta-se a necessidade de ressaltar aspectos referentes ao problema da pesquisa e aos seus objetivos propostos.

Desse modo, quanto a questão norteadora da pesquisa: **De que modo as estratégias de conversão de metadados digitais como o MARCXML, *Metadata Authority Description Schema* (MADS) e *Metadata Object Description Schema* (MODS) ao incorporarem atributos da descrição de autoridades de nomes e assuntos do formato MARC, potencializam o enriquecimento lógico-semântico a partir da estrutura conceitual dos FRBR?**Foi possível observar que a conjugação de esquemas de metadados e paradigmas da modelagem conceitual resultam na expressividade semântica dos pontos de acesso de um catálogo, o que permite explorar estratégias dinâmicas de conversão para o domínio bibliográfico independente do suporte, conforme verificado no alcance dos objetivos da pesquisa descritos nos parágrafos abaixo.

O alcance da resposta à questão acima nos conduziu aos objetivos propostos, a iniciar pela tarefa de **mapear os campos de autoridade de nome e assunto do formato MARC bem como as entidades definidas no FRBR**, o que viabilizou a verificação da relação exaustiva dos atributos de nome pessoal e assunto, provendo elementos estratégicos para o alcance do objetivo seguinte;

Ao correlacionar os campos mapeados com propriedades e atributos dos metadados digitais MARCXML, MADS e MODS da *Library of Congress*, foi constatada a flexibilidade notacional provida pela linguagem XML, propiciada pela inserção dos elementos extraídos do objetivo anterior em paralelo às relações internas providas pelos esquemas em estudo;

Por último, quanto à **identificação e a representação dos recursos catalogados em MARC21 nos metadados digitais MARCXML, MADS e MODS da *Library of Congress* partir do modelo entidade-relacionamento dos FRBR**, verificou-se que o modelo relacional adotado pelos requisitos funcionais tem o condão de estabelecer e ampliar a relação de atributos e o estabelecimento de vínculos inerentes

às autoridades e entidades, tornando a conversão de metadados no domínio bibliográfico uma tarefa capaz de ampliar o potencial descritivo com redução de perdas de informação.

À guisa de sugestão para pesquisas futuras, a despeito do paradoxo de se tratar de tema pouco explorado em língua portuguesa e, ao mesmo tempo, amplamente demandado como requisito para plataformas *web*, considera-se:

- 1) explorar modelos conceituais a partir dos diagramas constantes nos anexos da documentação dos FRAD e FRSAD nos estudos que tratem de autoridades de nomes e assuntos respectivamente;
- 2) desenvolver *frameworks* destinados à conversão de metadados para catálogos de autoridade, particularmente MADS e MODS;
- 3) criar estratégias de conversão para todas as entidades do domínio bibliográfico tratadas pelos FRBR com ênfase na integridade dos campos e subcampos MARC;
- 4) atualizar pesquisas sobre perdas de informação na conversão reversa de MARCXML, MADS e MODS para MARC;
- 5) aplicar propriedades dos esquemas MADS e MODS como vínculos às entidades dos grupos 2 e 3 dos FRBR.

Finalmente, infere-se que todas as linguagens citadas nesta pesquisa, além da anotação semântica que propiciam, são legíveis por homens e máquinas e, estas últimas, compreendem as ambiências multiplataformas tão recorrentes na atualidade, tais como *tablets*, *smartphones* e *laptops*. Nesse contexto, os novos paradigmas da catalogação, ao dedicarem atenção à *web*, tem enfatizado os tratamentos sintático e semântico do conteúdo informacional, aliado aos modelos conceituais que passam a figurar como parâmetros estratégicos capazes de operacionalizar conversões efetivas.

Isto posto, os achados da pesquisa abrem perspectivas de aprofundamento do tema e abordagens que possam considerar a relevância do paradigma relacional ao domínio bibliográfico.

REFERENCIAS

AALBERG, Trond; ZUMER, Maja. The value of MARC data, or, challenges of frbrisation. **Journal Of Documentation**, [s.l.], v. 6, n. 69, p.851-872, dez. 2012.

ALMEIDA, Maurício Barcellos; CENDÓN, Beatriz Valadares. Pesquisa sobre ferramentas de conversão de registros catalográficos padrão MARC para a linguagem XML. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. **Anais...** Belo Horizonte: UFMG, 2003.

ALVES, Rachel Cristina Vesú. **Metadados como elementos do processo de catalogação**. 2010. 134 f. Tese (Doutorado) - Doutorado em Ciência da Informação, UNESP, Marília, 2010.

ALVES, Rachel Cristina Vesú. **Web Semântica: uma análise focada no uso de metadados**. 2005. 182f. Dissertação (Mestrado) - Mestrado em Ciência da Informação, UNESP, Marília, 2005.

ASSUMPÇÃO, Fabrício Silva; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Metadata Authority Description Schema (MADS): uma alternativa à utilização do formato MARC21 para dados de autoridade. **Inf. Inf.**, Londrina, v. 18, n. 1, p.106-126, jan. 2013.

ASSUMPÇÃO, Fabrício Silva. **Conversão de registros em XML para MARC 21: um modelo baseado em folhas de estilo XSLT**. 2013. 137 f. Dissertação (Mestrado) - Mestrado em Ciência da Informação, UNESP, Marília, 2013.

BAGLEY, Philip. Extension of programming language concepts, Philadelphia, University City Science Center, 1969, p. 26. Disponível em: <<http://stevenbagley.net/blog/philip-r-bagley.html>> Acesso em 20 de nov. 2014.

BAKER, Thomas. Libraries, languages of description, and linked data: a Dublin Core perspective. **Library Hi Tech**, [s.l.], v. 30, n. 1, p.116-133, 2 mar. 2012.

BEZERRA, Darlene Alves. **O modelo FRBR e a busca de semântica na catalogação e recuperação de informações em ambientes digitais**. 2013. 187 f. Dissertação (Mestrado) - Curso de Mestrado em Ciência da Informação, Universidade Federal Fluminense, Niterói, 2013.

BUCKLAND, Michael K. Descrição e pesquisa: metadados como infra-estrutura. **BJIS**, Marília, v. 0, n. 0, p.3-15, jan./jul. 2006.

CARRASCO, Rafael C.; SERRANO, Aureo; CASTILLO-BUERGO, Reydi. A parser for authority control of author names in bibliographic records. **Information Processing And Management**, [s.l.], v. 1, n. 52, p.753-764, abr. 2016.

CASTRO, Fabiano Ferreira de. Functional requirements for bibliographic description in digital environments. **Transinformação**, Campinas, v. 2, n. 28, p.223-231, maio/jun. 2016.

CASTRO, Fabiano Ferreira de.; SIMIONATO, Ana Carolina; ZAFALON, Zaira Regina. Aspectos relacionais entre ontologia e metadados: considerações interdisciplinares. In: ENANCIB, 17., 2016, João Pessoa, PB. **Anais...** . Salvador , BA: ANCIB, 2016. p. 1 - 20.

CATARINO, Maria Elisabete; SOUZA, Terezinha Batista de. A representação descritiva no contexto da web semântica. **TransInformação**, Campinas, v.2, n.24, p.77-90, maio/ago., 2012

CHEN, Peter Pin-shan. The Entity-Relationship Model: toward a unified view of data. **ACM**, Massachusetts, v. 1, n. 1, p.9-36, mar. 1976.

COLE, Timothy W.. Qualified Dublin Core metadata for online journal articles. OCLC Systems & Services: **International digital library perspectives**, [s.l.], v. 18, n. 2, p.79-87, jun. 2002.

COYLE, Karen. RDA in RDF. **Library Technology Reports**, v. 5, n. 4, p.26-36, fev./mar. 2010.

CUNDIFF, Morgan V.. An introduction to the Metadata Encoding and Transmission Standard (METS). **Library Hi Tech**, [s.l.], v. 22, n. 1, p.52-64, mar. 2004.

DECOURSELLE, Joffrey et al. Open datasets for evaluating the interpretation of bibliographic records. **JCDL**, Newark, Nj, p.253-254, jun. 2016.

EL-SHERBINI, Magda; KLIM, George. Metadata and cataloging practices. **The Electronic Library**, [s.l.], v. 22, n. 3, p.238-248, jun. 2004.

FACHIN, Gleisy Regina Bories. **Ontologia de referência para periódico científico digital**. 2011. 407 f. Tese (Doutorado) - Curso de Doutorado em Ciência da Informação, Ufsc, Florianópolis, 2011.

FERREIRA, Jaider Andrade; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. **Inf. & Soc**, João Pessoa, v. 23, n. 2, p.13-23, maio/ago. 2013.

FETTERS, Linda K. **Handbook of indexing techniques**. Corpus Christ, TX: FimCoBooks, 2001. 128 p.

FLAMINO, Adriana Nascimento. **MARCXML**: um padrão de descrição para recursos informacionais em Open Archives. 2006. 166 f. Dissertação (Mestrado) - Curso de Mestrado em Ciência da Informação, UNESP, Marília, 2006.

FORD, Kevin. Library of Congress classification as linked data. **JLIS**, v. 4, n. 1, p.1-15, jan. 2013.

FURGERI, Sérgio. O papel das linguagens de marcação para a Ciência da Informação. **Transinformação**, Campinas, v. 3, n. 18, p.225-239, set. 2006.

FUSCO, Elvis. **Modelos conceituais de dados como parte do processo da catalogação**: perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. 2010. 251 f. Tese (Doutorado) - Curso de Doutorado em Ciência da Informação, Unesp, Marília, 2010.

GARTNER, Richard. An XML schema for enhancing the semantic interoperability of archival description. **Archival Science**, [s.l.], v. 15, n. 3, p.295-313, 1 ago. 2014.

GÓMEZ, Paola Andrea Picco. El cambio en la designación general de material [DGM] de las AACR a las RDA: definición de los atributos tipo de contenido, tipo de medio y tipo de soporte. **La Plata**, v. 1, n. 2, p.1-12, jul. 2012.

GONZALES, Brigid Mooney. The conversion of MARC metadata for online visual resource collections: a case study of tactics, challenges and results. **Library Philosophy And Practice**, Nebraska, p.1-66, fev. 2014.

GUENTHER, Rebecca. MADS: what it stands for: Metadata Authority Description Schema. **Computers In Libraries**, Washington, DC, v. 4, n. 27, p.1-14, abr. 2007.

GUENTHER, Rebeca; MCCALLUM, Sally. New metadata standards for digital resource: MODS and METS. **American Society For Information Science And Technology**, S.l, v. 29, n. 2, p.12-15, 2003.

GUENTHER, Rebeca. MODS: The Metadata Object Description Schema. **Libraries and the Academy**, Baltimore, MD, v. 3, n. 1, p. 137-150, jan. 2003.

GUENTHER, Rebeca. Using the Metadata Object Description Schema (MODS) for resource description: guidelines and applications. **Library Hi Tech**, [s.l.], v. 22, n. 1, p.89-98, mar. 2004.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Bibliographic Records**: final report. [S.l.]: IFLA, 2008.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **International Standard Bibliographic Description (ISBD)**. [S.l.]: IFLA, 2009.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Authority Data**: a conceptual model. [S.l.]: IFLA, 2013.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Functional Requirements for Subject Authority Data (FRSAD)**: a conceptual model. [S.l.]: IFLA, 2010.

JAKSIC, Mirjana. Mapping of bibliographical standards into XML. **Softw. Pract. Exper**, v. 34, n. 1, p.1051-1064, jan. 2004.

JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA. RDA: Resource Description and Access. [S.l.], 2010. Disponível em <<http://www.rdajsc.org/docs/rdabrochureJanuary2010.pdf>>. Acesso em: 31 maio 2012.

KEITH, Corey. Using XSLT to manipulate MARC metadata. **Library Hi Tech**, [s.l.], v. 22, n. 2, p.122-130, nov. 2003.

LAM, Ki-tat. XML and global name access control. OCLC Systems & Services: **International digital library perspectives**, [s.l.], v. 18, n. 2, p.88-96, jun. 2002.

LIBRARY OF CONGRESS. **METS**: Metadata Encoding & Transmission Standard. Washington, DF: LOC, 2016. Disponível em: <<http://www.loc.gov/standards/mets/>>. Acesso em: 06 abr. 2017.

LIBRARY OF CONGRESS. **MODS**: Metadata Object Description Schema. Washington, DF: LOC, 2016. Disponível em: <<http://www.loc.gov/standards/mods/>>. Acesso em: 08 abr. 2018.

LIBRARY OF CONGRESS. **MARCXML**: uses and features. Washington, DF: LOC, 2016. Disponível em: <<http://www.loc.gov/standards/marcxml/marcxml-overview.html>>. Acesso em: 06 abr. 2017.

LONG, Chris E.. Disambiguating the departed: Using the genealogist's tools to uniquely identify the long-dead and little-known. **LRTS**, Boulder, v. 4, n. 60, p.236-247, jun. 2016.

MACHADO, Denise Ramires *et al.* A cada autor, seu nome. A cada nome, seu documento.: política de controle de autoridades em ambiente de catalogação descentralizada e cooperativa. In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9., 2013, Rio de Janeiro. **Anais...** . Rio de Janeiro: Ufrgs, 2013. p. 1 - 19.

MACHADO, Raquel Bernadete; BRIGIDI, Fabiana Hennies. Controle de autoridade de nomes pessoais: relato de experiência do SIBI/UFSC. In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9., 2013, Rio de Janeiro. **Anais...** . Rio de Janeiro: Ufrgs, 2013. p. 1 - 16.

MADS. Metadata Authority Description Schema. Washington DC: Library of Congress, 2016.

MAIMONE, Giovana Deliberali; SILVEIRA, Naira Christofolletti; TÁLAMO, Maria de Fátima Gonçalves. Moreira. Reflexões acerca das relações entre representação temática e descritiva. **Inf. & Soc.**, João Pessoa, v. 21, n. 1, p.27-35, jan. 2011.

MARC IN XML. Marc Standards. Washington DC: Library of Congress, 2008.

MARC TO MODS. Metadata Object Description Schema. Washington DC: Library of Congress, 2014.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Técnicas de pesquisa**. 7 ed. São Paulo: Atlas, 2016. 277p.

MCCALLUM, Sally H.. An introduction to the Metadata Object Description Schema (MODS). **Library Hi Tech**, [s.l.], v. 22, n. 1, p.82-88, mar. 2004.

MCDONOUGH, Jerome P. METS: standardized encoding for digital library objects. **International Journal On Digital Libraries**, [s.l.], v. 6, n. 2, p.148-158, 1 fev. 2006.

MELO, Maria Antônia Fonseca. **Contribuições das abordagens positivista e pragmática do estudo do conceito para o modelo conceitual FRISAD**. 2013. 133 f. Dissertação (Mestrado) - Pós-Graduação em Ciência da Informação, UnB, Brasília, 2013.

MEY, Eliane Serrão Alves; GRAU, Isabel Arino; BIAR, Fernanda Salgado. Resource Description and Access (RDA): prós e contras. **R. Ibero-amer.**, Brasília, v. 7, n. 1, p.43-52, jan./jul. 2014.

MOULAISON, Heather Lea. The expansion of the personal name authority record under resource description and access: current status and quality considerations. **IFLA Journal**, London, v. 1, n. 41, p.13-24, fev. 2015.

MORENO, Fernanda Passini; BRASCHER, Marisa. MARC, MARCXML e FRBR: relações encontradas na literatura. **Inf. & Soc.**, João Pessoa, v. 3, n. 17, p.13-25, set./dez. 2007.

OCHANDO, Manuel Blázquez. Redifusión de catálogos bibliográficos en MARC-XML. **Investigación Bibliotecológica**, México, DF, v. 28, n. 62, p.53-65, jan./abr. 2014.

RALSON, Katherine N.. METS: Flexibility v. Interoperability. **Library Philosophy And Practice**, Lincoln Ne, v. 1, n. 1, p.1-26, set. 2013.

RICHARDSON, Roberto Jarry. **Pesquisa social: métodos e técnicas**. 3 ed. rev. e ampl. São Paulo: Atlas, 2011. 334 p.

RILEY, Jenn. **Glossary of metadata standards**. Indianapolis: IUL, 2010.

SALGADO, Denise Mansera. **O controle de autoridade sob a norma RDA: análise da aplicação e implicações na construção de registros de autoridade**. 2015. 173 f. Dissertação (Mestrado) - Curso de Mestrado em Ciência da Informação, USP, São Paulo, 2015.

SCHMIDT, Nikita; PATEL, Ahmed. MARC record conversion: a generalised approach. **Computer Standards & Interfaces**, Dublin, v. 1, n. 21, p.287-297, mar. 1999.

SENGUPTA, Arijit. On the feasibility of using conceptual modeling constructs for the design and analysis of XML data. **Data & Knowledge Engineering**, [s.l.], v. 72, p.219-238, fev. 2012.

SILVEIRA, Naira Christofolletti; TÁLAMO, Maria de Fátima Gonçalves Moreira. Os FRBR e a escolha do ponto de acesso principal. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 2, n. 14, p.108-120, maio/ago. 2009.

SIQUEIRA, Marcos Antonio. **XML na Ciência da Informação**: uma análise do MARC 21. 2003. 133 f. Dissertação (Mestrado) - Curso de Mestrado em Ciência da Informação, UNESP, Marília, 2003.

SNYMAN, Mmm; VAN RENSBURG, M Jansen. Reengineering name authority control. **The Electronic Library**, [s.l.], v. 17, n. 5, p.313-322, out. 1999.

SOUZA, Elisabete Gonçalves de; BEZERRA, Darlene Alves; COSTA, Wellington Freire Cunha. Descrição de recursos em uma estrutura de metadados pautada no modelo FRBR. **em Questão**, Porto Alegre, RS, v. 22, n. 1, p.113-136, jan. 2016.

TAHERI, Sayyed Mahdi; HARIRI, Nadjla; FATTAHI, Sayyed Ramatollah. Does discarding XML declarations and changing file extensions improve the indexability and visibility of metadata tag names in web search engines? **Journal Of Information Science**, v. 40, n. 6, p.796-805, jan. 2014.

TOSAKA, Yuji; PARK, Jung-ran. RDA: Resource description & access—a survey of the current state of the art. **Journal of the American Society for Information Science and Technology**, [s.l.], v. 4, n. 64, p.651-662, set. 2013.

TRIVIÑOS, Augusto Nivaldo Silva. **Introdução à pesquisa em Ciências Sociais**: a pesquisa qualitativa em educação. São Paulo: Atlas, 2013. 175p.

UNDERSTANDING Metadata. Bethesda, Md: NISO Press, 2004.

UNIFIED MODELLING LANGUAGE. UML Superstructure Specification. [S.l.]: OMG, 2010.

VIDAKOVIĆ, Jovana; RACKOVIĆ, Miloš. Generating content and display of library catalogue cards using XML technology. **Software: Practice and Experience**, [s.l.], v. 36, n. 5, p.513-524, 2006.

WAGNER, Lindsey. METS: a survey of recent literature and applications. **Library Philosophy and Practice**, Lincoln Ne, v. 0, n. 0, p.1-8, jan. 2011.

W3C Information and knowledge domain. **Extensible Markup Language (XML)**. 2016. Disponível em: <<https://www.w3.org/XML/>>. Acesso em: 20 mar. 2017.

W3C. **XML Schema Part 1**: structures second edition. 2004. Disponível em: <<https://www.w3.org/TR/xmlschema-1/>>. Acesso em: 25 abr. 2017.

W3C. **URIs, URLs, and URNs**: Clarifications and Recommendations 1.0. Disponível em: <<https://www.w3.org/TR/uri-clarification/>>. Acesso em: 23 mar. 2018.

YEE, Martha M.. New perspectives on the shared cataloging environment and a MARC 21 shopping list. **LRTS**, [s.l.], v. 3, n. 48, p.165-178, set. 2004.

ZAFALON, Zaira Regina. Sintaxe e semântica de registros bibliográficos: princípios para a conversão de registros analógicos para o Formato MARC21 Bibliográfico: Scan for MARC. **TPBCI**. Salvador, BA, v. 5, n. 1, p.1-23, 2012.

ZAFALON, Zaira Regina. **Scan for MARC**: princípios sintáticos e semânticos de registros bibliográficos aplicados à conversão de dados analógicos para o formato MARC 21 bibliográfico. 2012. 170 f. Tese (Doutorado) - Curso de Doutorado em Ciência da Informação, UNESP, Marília, 2012.

ZAVALINA, Oksana L.; SHAKERI, Shadi; KIZHAKKETHIL, Priya. Metadata change in traditional library collections and digital repositories: exploratory comparative analysis. **ASIST**, St. Louis, MO, p.1-5, 2015.

ZENG, Marcia Lei. Metadata elements for object description and representation: a case report from a digitized historical fashion collection project. **Journal Of The American Society For Information Science**, Kent, Oh, v. 13, n. 50, p.1193-1208, jan. 1999.