

unesp  **UNIVERSIDADE ESTADUAL PAULISTA**
"JÚLIO DE MESQUITA FILHO"
Faculdade de Filosofia e Ciências,
Campus de Marília – SP
Programa de Pós-Graduação em Ciência da Informação

FABIANO FERREIRA DE CASTRO

**ELEMENTOS DE INTEROPERABILIDADE NA CATALOGAÇÃO DESCRITIVA:
configurações contemporâneas para a modelagem de ambientes informacionais digitais**



Marília
2012

FABIANO FERREIRA DE CASTRO



**ELEMENTOS DE INTEROPERABILIDADE NA CATALOGAÇÃO DESCRITIVA:
configurações contemporâneas para a modelagem de ambientes informacionais digitais**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação, como parte das exigências para a obtenção do título de Doutor em Ciência da Informação, da Faculdade de Filosofia e Ciências – Universidade Estadual Paulista - UNESP, Campus de Marília. **Área:** Informação, Tecnologia e Conhecimento.

Linha de Pesquisa: Informação e Tecnologia

Orientadora: Profa. Dra. Plácida Leopoldina Ventura Amorim da Costa Santos

Marília
2012

C353e Castro, Fabiano Ferreira de.
Elementos de interoperabilidade na catalogação descritiva:
configurações contemporâneas para a modelagem de ambientes
informacionais digitais / Fabiano Ferreira de Castro. - - Marília, F.
F. de Castro, 2012.
202 f.: il. ; 30 cm

Tese (Doutorado em Ciência da Informação) – Faculdade de
Filosofia e Ciências – Universidade Estadual Paulista, 2012.

Bibliografia: f. 192-202.

Orientadora: Profa. Dra. Plácida Leopoldina Ventura Amorim
da Costa Santos.

1. Catalogação descritiva. 2. Ambientes informacionais digitais. 3.
Metadados. 4. Ontologia para descrição bibliográfica. 5.
Interoperabilidade. I. Autor. II. Título.

CDD 025.316

FABIANO FERREIRA DE CASTRO

ELEMENTOS DE INTEROPERABILIDADE NA CATALOGAÇÃO DESCRITIVA: configurações contemporâneas para a modelagem de ambientes informacionais digitais

BANCA EXAMINADORA:

Nome: Profa. Dra. Plácida Leopoldina Ventura Amorim da Costa Santos.

Titulação: Profa. do Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências – Universidade Estadual Paulista – UNESP, Campus de Marília. (Orientadora – Presidente da Banca).

Nome: Profa. Dra. Lídia Alvarenga.

Titulação: Profa. do Programa de Pós-Graduação em Ciência da Informação – Escola de Ciência da Informação – Universidade Federal de Minas Gerais – UFMG, Campus de Belo Horizonte.

Nome: Dr. Rogério Aparecido Sá Ramalho.

Titulação: Prof. do curso de Biblioteconomia e Ciência da Informação – Centro de Educação e Ciências Humanas – Universidade Federal de São Carlos – UFSCar, Campus de São Carlos.

Nome: Dr. Ricardo César Gonçalves de Sant’Ana.

Titulação: Prof. do Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências – Universidade Estadual Paulista – UNESP, Campus de Marília.

Nome: Dra. Silvana Aparecida B. Gregorio Vidotti.

Titulação: Profa. do Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências – Universidade Estadual Paulista – UNESP, Campus de Marília.

Data da defesa: 05/03/2012.

Local: Universidade Estadual Paulista – UNESP – Campus de Marília.

Dedico este trabalho à Deus e aos catalogadores contemporâneos.

AGRADECIMENTOS

Inicio meus agradecimentos, reconhecendo que só estou aqui e cheguei nesse estágio graças ao meu bom e amado Senhor Jesus Cristo, como diz a canção, [...] *eu posso ir, muito além de onde estou, vou nas asas do Senhor, o teu amor é o que me conduz* [...].

Agradeço aos meus familiares, meu pai Alaor, minha mãe Ana, minha irmã Tatiane e meu irmão Júnior, sem os quais não estaria concluindo mais esse estágio na minha vida.

À minha amada e querida Paula, por toda sua paciência comigo nesses momentos de ausência, com o seu jeito terno e carinhoso, sempre me acolhe e cuida de mim, pelo apoio e segurança incondicionais. Meu ponto de equilíbrio, cúmplice e amiga. Você é parte fundamental dessa nossa conquista. Te amo para sempre minha linda!

Aos meus sogros Aparecida (Tuta) e Moacir por todo o carinho, apoio e acolhimento na minha vida, pela amizade e confiança cativadas. Minha nova família.

À família Figueiredo, Carlão, Mi, Lucas, Helen e Thiago, que se tornaram também minha família, me acolhendo e propiciando todas as condições necessárias para meu bem-estar.

Agradeço em especial, minha orientadora Profa. Dra. Plácida, por todos os seus ensinamentos ao longo desses anos, por seu conhecimento e competência ímpares, por toda sua paciência, e por acreditar em mim, fico muito feliz pelo profissional que tenho me tornado, pois isso não seria possível sem a sua determinação e comprometimento na minha formação. Muito mais do que uma orientadora, ela é uma grande mãe e amiga e sempre têm as palavras certas na hora e nas mais diversificadas situações. Foi muito bom ter sido seu orientando estes anos, mas acredito que sua orientação se perpetuará por toda minha vida profissional e acadêmica.

Aos professores membros da banca examinadora por toda contribuição e apontamentos valiosos à pesquisa.

Aos professores do Programa de Pós-Graduação em Ciência da Informação da UNESP, pela formação profissional e acadêmica.

Aos meus colegas membros do Grupo de Pesquisa Novas Tecnologias em Informação, por toda oportunidade de reflexão e crescimento científico.

Aos meus queridos amigos que me desculpem, mas não vou me atrever mencioná-los, correndo o perigo de esquecer algum nome, vocês sabem que são peças-chave na minha vida, e agradeço por toda amizade dedicada, alguns mais recentes, outros há uma vida inteira.

Aos meus colegas docentes do curso de Biblioteconomia da Universidade Federal de Sergipe (UFS), Martha, Glêyse, Sérgio, Nilton e de modo especial, minha amiga Valéria Bari, por todo apoio, compreensão e generosidade. Obrigado por tudo!

Aos meus alunos e orientandos do curso de Biblioteconomia da Universidade Federal de Sergipe (UFS), por me oportunizarem a vivência da tríade (ensino, pesquisa e extensão), na universidade. Minha motivação enquanto professor é porque vocês existem e apostaram suas “fichas” (não catalográficas) numa carreira promissora.

À Universidade Federal de Sergipe (UFS) por me liberar e me apoiar no meu doutoramento.

À Fundação de Amparo e Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro concedido nos primeiros anos do doutorado.

A todos aqueles que passaram por minha vida durante esse período e que deixaram suas marcas e perfumes, o meu muito obrigado!

"Não há ensino sem pesquisa e pesquisa sem ensino. Enquanto ensino continuo buscando, reprocurando. Ensino porque busco, porque indaguei, porque indago e me indago. Pesquiso para constatar, constatando intervenho, intervindo educo e me educo. Pesquiso para conhecer o que ainda não conheço e comunicar ou anunciar a verdade".

Paulo Freire

"É do buscar e não do achar que nasce o que eu não conhecia".

Clarice Lispector

RESUMO

O uso intensivo das Tecnologias da Informação e Comunicação (TIC) gera o desenvolvimento de uma gama de padrões de metadados, instrumentos fundamentais nos ambientes de manipulação de dados bibliográficos. Verifica-se, atualmente, uma tendência para a disponibilização de recursos e conteúdos informacionais em formato digital e a não utilização de padrões adequados para representar e descrever estes recursos, considerado alvo de preocupação da comunidade científica. A tese levantada indica que as metodologias da Catalogação Descritiva e uma ontologia para a descrição bibliográfica explicitada nos códigos e padrões de metadados favorecem a recuperação da informação, a partir da modelagem do banco de dados e da construção de formas de representação dos recursos informacionais, possibilitando a interoperabilidade em ambientes digitais. A proposta de pesquisa está pautada em verificar na literatura científica a aplicabilidade e a funcionalidade das tecnologias e dos instrumentos que estão disponíveis atualmente, para o tratamento descritivo das informações em ambientes informacionais digitais. Objetiva-se, a partir dos novos cenários da Catalogação Descritiva identificar e analisar a iniciativa MarcOnt, suas similaridades e diferenças com o padrão de metadados MARC 21, da área de Biblioteconomia, na tentativa de compará-los e articulá-los com as ferramentas para a padronização de descrição dos aspectos de forma e de conteúdo de recursos bibliográficos, componentes do conhecimento da Ciência da Informação. Com uma metodologia de análise exploratória e descritiva do tema, e fundamentada no Perspectivismo (PETERSON, 1996) será possível uma compreensão do MarcOnt e as tecnologias que estão abarcadas no âmbito das bibliotecas digitais para a construção *Web Semântica*. Como resultado dessa investigação é apresentado uma proposta de requisitos funcionais para a modelagem de ambientes informacionais digitais que garantem a interoperabilidade no domínio bibliográfico. Acredita-se que as metodologias da Catalogação Descritiva e uma ontologia para a descrição bibliográfica, explicitada nas regras e códigos de catalogação, e nos padrões de metadados redesenham a construção de novos ambientes informacionais digitais melhor estruturados, para a recuperação da informação e no estabelecimento efetivo da interoperabilidade.

Palavras-chave: Catalogação Descritiva. Modelagem de ambientes informacionais digitais. Metadados. Interoperabilidade. Ontologia para descrição bibliográfica. Recuperação da informação. Informação e Tecnologia.

Interoperability's elements in the Descriptive Cataloging: contemporary settings for modeling of digital informational environments

ABSTRACT

The intensive use of Information and Communication Technologies (ICT) generates the development of a various possibility of metadata standards, essential in environments manipulation of bibliographic data. Actually, there is a tendency for the availability of resources and informational content in digital format and not use of appropriate standards to represent and describe these features, considered the subject of concern for the scientific community. The thesis shows that methodologies of Descriptive Cataloging and bibliographic description of an ontology is explicit in the code and metadata standards favor of retrieval information from the database modeling and construction of forms of representation of information resources, enabling interoperability in digital environments. This proposal verifies in the scientific literature the applicability and functionality of technologies and mechanisms that are available currently as the information descriptive treatment in digital informational environments. Objective is, as new scenarios of Descriptive Cataloging identify and analyze for the initiative MarcOnt, its similarities and differences regarding the MARC 21 metadata standard of the Librarianship area, in an attempt to compare and link them with tools that standardize and describe form and content aspects of bibliographic resources, knowledge components of Information Science. With a methodology exploratory and descriptive analysis of the subject, based on the Perspectivism (PETERSON, 1996) can be MarcOnt and an understanding of the technologies that are involved within the scope of digital libraries for building Semantic Web. As a result of this investigation is presented a proposal for modeling a functional requirements of digital informational environments that ensure interoperability in the bibliographic domain. It is believed that the methods of descriptive cataloging and bibliographic description of an ontology, explicit in the rules and codes of cataloging and metadata standards in the redesigning the building of new digital informational environments better structured to ensure the retrieval information and the establishment of effective interoperability.

Keywords: Descriptive cataloging. Modeling of digital informational environments. Metadata. Interoperability. Bibliographic description of an ontology. Information retrieval. Information and Technology.

LISTA DE ILUSTRAÇÕES

FIGURA 1: Caso de heterogeneidade.....	37
FIGURA 2: Abordagens e tipologias de ontologias segundo Gruber et al. (2007).....	42
FIGURA 3: Arquitetura da <i>Web Semântica</i> comentada pelo autor.....	48
FIGURA 4: Etiqueta título do formato MARC 21	58
FIGURA 5: Estrutura do formato MARC 21	65
FIGURA 6: Exemplo de mapeamento entre os elementos de DC e MARC 21	70
FIGURA 7: Estrutura geral de classes da ontologia MarcOnt.	81
FIGURA 8: Recurso informacional livro com seus atributos	83
FIGURA 9: Descrição de um recurso bibliográfico utilizando a ontologia MarcOnt.....	85
FIGURA 10: Arquitetura de serviços de mediação MarcOnt	106
FIGURA 11: RDF como <i>core</i> para descrições bibliográficas semânticas	108
FIGURA 12: Perda de dados sem o uso de RDF.	110
FIGURA 13: Adoção de RDF para interoperabilidade semântica	111
FIGURA 14: Arquitetura RDF simples.....	135
FIGURA 15: Declaração RDF utilizando a sintaxe XML	136
FIGURA 16: Criação de propriedades RDA para entidades FRBR.....	147
FIGURA 17: Propriedade livro em RDA.....	148
FIGURA 18: Propriedades e subpropriedade RDA e FRBR.	148
FIGURA 19: Criação do elemento Extent.....	149
FIGURA 20: Caracterizações não incluídas no modelo RDA e RDF.....	152
FIGURA 21: Atributo Creator.....	152
FIGURA 22: Estrutura de declarações pré-coordenadas.....	155
FIGURA 23: Declaração de publicação	155
FIGURA 24: Declaração pré-coordenada como <i>string</i> : estrutura.....	156
FIGURA 25: Instância 1: Lugar de publicação	157
FIGURA 26: Instância 2: Lugar de publicação.....	158
FIGURA 27: Diretrizes funcionais para a interoperabilidade	175
FIGURA 28: Proposta de modelagem para interoperabilidade no domínio bibliográfico.....	177
FIGURA 29: Arquitetura para a Representação e a Descrição Bibliográfica semântica e níveis de interoperabilidade	178

LISTA DE QUADROS

Quadro 1: Pesquisa em fontes primárias e secundárias	24
Quadro 2 : Analogia da ontologia com a Catalogação Descritiva.....	40
Quadro 3: Classificação de ontologias	41
Quadro 4 : Sistematização das investigações contemporâneas sobre ontologias sob o prisma da Ciência da Informação e Ciência da Computação	47
Quadro 5: Principais propriedades da ontologia MarcOnt	84
Quadro 6: Conceitos dos recursos informacionais em MarcOnt.....	113
Quadro 7: Conceitos dos recursos informacionais em AACR2	114
Quadro 8: Conceitos dos recursos informacionais em RDA.....	115
Quadro 9: Principais conceitos dos recursos informacionais encontrados em MARC 21	116
Quadro 10: Sistematização conceitual dos principais recursos de MarcOnt em relação ao AACR2 e RDA.....	117
Quadro 11: Diferenças terminológicas nos modelos de dados.....	170
Quadro 12: Possíveis realocações de funções atuais para a <i>Web Semântica</i>	172
Quadro 13: Recomendações funcionais para bibliotecas digitais e <i>Web Semântica</i>	174
Quadro 14: Explicitação de ontologia para os elementos autor e título em AACR2	179
Quadro 15: Explicitação de ontologia para elementos autor e título em MARC 21	180
Quadro 16: Regras de conteúdo definidas por AACR2 para elementos autor e título em MARC 21	181

LISTA DE SIGLAS

AACR - Anglo-American Cataloging Rules
CDD – Classificação Decimal de Dewey
CDU – Classificação Decimal Universal
CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico
DC - Dublin Core
DCMI - Dublin Core Metadata Initiative
DERI – Digital Enterprise Research Institute
DER - Diagrama de Entidade Relacionamento
DTD - Document Type Definition
ER – Entidade relacionamento
FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo
FOAF - Friend Of A Friend
FRBR - Functional Requirements of Bibliographic Records
FRAD - Functional Requirements for Authority Data
FRSAD - Functional Requirements for Subject Authority Data
GUT - Gdańsk University of Technology
IFLA - International Federation of Library Association and Institutions
ISBD - International Standard Bibliographic Description
JSC - Joint Steering Committee for Development of RDA
LC - Library of Congress
LCSH - Library of Congress Subject Headings
MARC - Machine Readable Cataloging
METS - Metadata Encoding and Transmission Standard
MODS - Metadata Object Description Schema
NACO - Name Authority Cooperative Program
NISO - National Information Standards Organization
OWL – Ontology Web Language
RDA - Resource Description and Access
RDF - Resource Description Framework
RDFS - Resource Description Framework Schema
SGML - Standard Generalized Markup Language
SKOS - Simple Knowledge Organization System

SES - Syntax Encoding Scheme

SRW - Search Retrieval Web Service

TDI – Tratamento Descritivo da Informação

TIC - Tecnologias da Informação e Comunicação

UNESP - Universidade Estadual Paulista “Júlio de Mesquita Filho”

URI - Uniform Resource Identifier

URL - Uniform Resource Locator

VRA - Visual Resources Associations

W3C - World Wide Web Consortium

XHTML - eXtensible HyperText Markup Language

XML - Extensible Markup Language

XSL - Extensible Stylesheet Language

SUMÁRIO

1 INTRODUÇÃO	18
1.1 Problema de pesquisa.....	20
1.2 Proposição e Tese.....	21
1.3 Objetivos.....	22
1.3.1 Objetivo Gerais	22
1.3.2 Objetivos Específicos.....	22
1.4 Metodologia.....	23
1.4.1 Procedimentos metodológicos.....	24
1.4.2 Forma de análise dos resultados.....	27
1.5 Justificativa	28
1.6 Estrutura da Tese	31
2 ONTOLOGIAS E METADADOS: aproximações sinérgicas para o estabelecimento da interoperabilidade em ambientes digitais	34
2.1 Ontologias: algumas considerações.....	34
2.2 Metadados: exigência para a modelagem de ambientes informacionais digitais ...	55
2.2.1 Padrão de metadados MARC 21: complexidade na representação e descrição bibliográfica	60
2.3 Crosswalk: mapeamento entre esquemas de metadados.....	67
2.4 Arquitetura de metadados: requisitos para a interoperabilidade	71
2.5 Integração estratégica entre ontologias e metadados.....	72
3 MARCONT: um diálogo possível com MARC 21?.....	75
3.1 Gênese da iniciativa MarcOnt.....	75
3.2 Estrutura do MarcOnt: classes e propriedades.....	81
3.3 Interoperabilidade entre padrões de metadados: o delineamento da arquitetura MarcOnt.....	106
3.3.1 Análise da arquitetura MarcOnt à luz da Catalogação Descritiva e dos princípios de interoperabilidade semântica	108
3.4 Análise comparativa conceitual entre as principais terminologias (classes) do MarcOnt à luz dos esquemas de descrição da Biblioteconomia e Ciência da Informação (AACR2 e RDA)	112

3.5 Considerações e apontamentos sobre o MarcOnt	120
4 RESOURCE DESCRIPTION FRAMEWORK (RDF): promessa para interoperabilidade em ambientes informacionais digitais.....	124
4.1 A visão de interoperabilidade no domínio bibliográfico.....	124
4.2 Resource Description Framework (RDF): o <i>core</i> para descrições bibliográficas	131
4.3 RDA: proposta de um código internacional de catalogação aplicado em ambientes digitais.....	140
4.4 O papel do RDF nos esquemas e modelos bibliográficos.....	142
5 CATALOGAÇÃO DESCRITIVA NO SÉCULO XXI: proposta de um modelo para a descrição bibliográfica semântica em ambientes informacionais digitais interoperáveis	161
5.1 A Catalogação Descritiva e sua repercussão nos ambientes informacionais digitais	161
5.2 Modelagem de dados para construção de ambientes digitais.....	165
5.3 Proposta de um modelo para a descrição bibliográfica semântica em ambiente digital	171
6 CONSIDERAÇÕES FINAIS.....	185
REFERÊNCIAS	192



CAPÍTULO 1



INTRODUÇÃO

1 INTRODUÇÃO

Não se pode procurar nem o que se conhece, nem o que não se conhece. O que se conhece porque, conhecendo-o, não se tem necessidade de procurá-lo; o que não se conhece porque não se sabe o que se deve procurar.
Platão

Inserida na linha de pesquisa Informação e Tecnologia do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP - campus de Marília, esta tese tem como temática a *Catálogo Descritiva no âmbito digital*, cuja delimitação é o estudo das ferramentas tecnológicas no interior da Representação Descritiva, na esfera das bibliotecas digitais e da *Web* de um modo geral, a partir da análise dos códigos e esquemas de descrição bibliográfica da área de Biblioteconomia e dos padrões de metadados, na perspectiva da Ciência da Informação.

Vale ressaltar que a presente pesquisa faz parte do projeto integrado financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e intitulado “Catálogo e recursos informacionais digitais: intersemioses, metadados e interoperabilidade em ambientes informacionais”, sob a coordenação da professora Dra. Plácida Leopoldina Ventura Amorim da Costa Santos, junto ao Grupo de Pesquisa Novas Tecnologias em Informação (GPNTI)¹ e vinculada à linha de pesquisa Ambientes informacionais digitais. Salienta-se que esta pesquisa foi financiada, *a priori*, pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), outorgada no processo 2008/53664-5, referente ao período de (22.05.09 a 10.07.10).

¹ Informações sobre o grupo de pesquisa podem ser encontradas no *site*: Disponível em: <<http://dgp.cnpq.br/buscaoperacional/detalhegrupo.jsp?grupo=0330607CJEMJIF>>. Acesso em: 01 fev. 2012.

É notório que nossa sociedade é marcada por constantes mudanças vivenciadas nos mais variados segmentos, uma delas de cunho tecnológico, no contexto que se conhece hoje por Sociedade da Informação² ou do Conhecimento. Percebe-se, que nos últimos, tempos houve um aumento desordenado e caótico na quantidade de informações produzidas e disponibilizadas em meio digital, requerendo uma mudança e um repensar nas formas de armazenamento, de representação, de descrição e de preservação dos recursos informacionais digitais.

As bibliotecas, em diferentes épocas, sempre se utilizaram das tecnologias disponíveis acompanhando, dessa forma, o desenvolvimento de paradigmas tecnológicos. Com relação às Tecnologias da Informação e Comunicação (TIC), desde a iniciativa de Vannevar Bush em 1945 com a criação do *Memex*, é possível observar o trabalho árduo de diferenciados profissionais no estabelecimento de mecanismos que pudessem agilizar os processos de tratamento e armazenamento das informações, a fim de garantir a sua integridade e recuperação.

A evolução das bibliotecas tem merecido grande destaque, ao longo do século XX, no que se refere ao desenvolvimento e uso de tecnologias, principalmente as TIC, pois potencializaram seus serviços com o oferecimento de novos recursos de acesso, com a utilização de formatos de intercâmbio de dados na Catalogação Descritiva, no uso de ontologias, na orientação para a modelagem de catálogos e no processo de busca e recuperação da informação.

Nessa perspectiva, tem-se o papel das bibliotecas digitais que atuam como ambientes estruturados na Internet e que passam a ser totalmente dependentes de tecnologias e precisam se adaptar às novas perspectivas tecnológicas na constituição de ambientes informacionais digitais e novos formatos de armazenamento, de descrição e de representação das informações, na garantia de estabelecer a interoperabilidade entre sistemas heterogêneos.

² De acordo com Mattelart (2002, p. 08) “Essa noção de sociedade da informação se formaliza na sequência das máquinas inteligentes criadas ao longo da Segunda Guerra Mundial”. Ela entra nas referências acadêmicas, políticas e econômicas a partir do final dos anos 1960. Durante a década seguinte, a fábrica que produz o imaginário em torno da nova “era da informação” já funciona a pleno vapor. Os neologismos lançados na época para designar a nova sociedade só mostrarão seu verdadeiro sentido geopolítico às vésperas do terceiro milênio com o que se convencionou chamar de “revolução da informação” e com a emergência da Internet como nova rede de acesso público.

1.1 Problema de pesquisa

Atualmente, no ambiente informacional digital, onde vários tipos de recursos coexistem com formatos e padrões de metadados heterogêneos, muitos esforços têm sido feitos para alcançar a interoperabilidade, a fim de utilizar vários padrões de metadados e de reutilizar os registros de metadados, através do desenvolvimento de estratégias, que vão desde os simples mapeamentos entre elementos de metadados à modelagem estrutural complexa.

Entende-se por interoperabilidade a capacidade de sistemas múltiplos com diferentes *hardwares* e *softwares*, plataformas, estruturas de dados e interfaces intercambiarem dados com o mínimo de perda de conteúdo e funcionalidade. Utilizando esquemas de metadados definidos, compartilhando transferência entre protocolos³, e *crosswalks*⁴ entre esquemas, os recursos na rede podem ser buscados mais amplamente. (ZENG; QIN, 2008).

Dessa forma, há uma preocupação da comunidade científica, no que tange à busca por caminhos e soluções para as questões de integração, interoperabilidade e entendimento semântico sobre os recursos informacionais disponíveis hoje nos ambientes informacionais digitais, de forma a possibilitar o acesso, o uso e o (re) uso das informações de forma mais efetiva para o usuário final.

Com o uso intensivo das Tecnologias da Informação e Comunicação (TIC) gera-se uma gama de padrões de metadados, instrumentos fundamentais nos ambientes de manipulação de dados bibliográficos.

Assim, é possível identificar que há uma tendência para a disponibilização de recursos e conteúdos informacionais em formato digital e a não utilização de padrões adequados para representar e descrever estes recursos. Elucida-se que o tratamento dado ao recurso informacional requer uma descrição de forma e de conteúdo legível por máquinas, com resultados compreensíveis aos humanos, e, que ainda, atenda aos requisitos de

³ O protocolo serve para acessar catálogos de outras instituições, independentemente do sistema utilizado, promove o acesso simultâneo a catálogos, compartilha registros bibliográficos e possui interface única para diferentes fontes.

⁴ Ferramenta utilizada para o mapeamento entre padrões de metadados heterogêneos.

interoperabilidade entre os ambientes informacionais, apontados pela literatura e focados nesta tese, enquanto problema da pesquisa.

Em busca de soluções para o problema analisa-se nesta tese a proposta de Kruk, Synak e Zimmermann (2005b), apontada nos princípios norteadores do *MarcOnt Initiative*, apresentando camadas estruturantes e delineadoras da *Web Semântica*⁵, tais como as ontologias e os esquemas de metadados, contemplando os objetivos das bibliotecas digitais e que propiciam meios mais adequados de representar, de organizar e permitir os relacionamentos entre os recursos informacionais digitais. E ainda analisa-se a ontologia para a descrição bibliográfica dos códigos de catalogação e padrões de metadados.

1.2 Proposição e Tese



Nesse contexto, a presente pesquisa pretende verificar na literatura científica a aplicabilidade e a funcionalidade das tecnologias e das ferramentas que estão disponíveis atualmente, para o Tratamento Descritivo das Informações (TDI), respaldadas nos instrumentos e métodos de descrição da Biblioteconomia, no âmbito digital e que podem ainda propiciar a interoperabilidade.

A tese defendida é a de que as metodologias, os instrumentos e os processos da Catalogação Descritiva, auxiliadas por uma ontologia para a descrição bibliográfica explicitada nos códigos e padrões de metadados favorecem a recuperação da informação, a partir da modelagem do banco de dados (catálogo bibliográfico) e da construção de formas de representação de recursos bibliográficos, na garantia da interoperabilidade em ambientes informacionais digitais.

⁵ De acordo com Santos e Alves (2009) a *Web Semântica* é a denominação dada para a estrutura tecnológica criada para estabelecer maior nível semântico aos dados; seu funcionamento envolve a implantação de ferramentas tecnológicas e métodos de representação da informação para o estabelecimento de raciocínio sobre os dados e, a partir daí, facilitar a recuperação e a construção de ambientes inteligentes.

1.3 Objetivos

A Ciência da Informação como uma área que tem como objeto de estudo a informação, desde a sua geração até o seu uso, pautada nos processos de representação, descrição, organização, acesso e recuperação de recursos informacionais, busca identificar metodologias e ferramentas tecnológicas para a construção de formas de representação de recursos bibliográficos em ambientes informacionais digitais. Dessa forma, esta investigação tem como objetivos:

1.3.1 Objetivos Gerais

Com os novos cenários da Catalogação Descritiva, marcados com a modelagem conceitual de catálogos com os Requisitos Funcionais para Registros Bibliográficos (FRBR), e a proposta do novo código de catalogação *Resource Description and Access* (RDA) e caracterizados por novas perspectivas tecnológicas, conhecidas hoje por *Web Semântica*, *Web 2.0*⁶ e *Web 3.0*⁷, a pesquisa tem por objetivo geral o estudo das ferramentas e dos métodos para o Tratamento Descritivo da Informação (TDI), em especial, para a representação e a descrição (forma e conteúdo) de recursos bibliográficos, no contexto digital, visando à apresentação de um *corpus* teórico.

1.3.2 Objetivos Específicos

Nesse sentido, os objetivos específicos, elencados para uma investigação mais sistemática, são os seguintes:

⁶ A *Web 2.0* se destaca pelo ambiente colaborativo e de interação para a construção e o compartilhamento do conhecimento. A sinergia criada por meio dessa colaboração e interação dos indivíduos acelera o processo de socialização do conhecimento, em espaços ou ambientes mais interativos e participativos. Os autores afirmam que a *Web 2.0* se constitui como: “*um novo espaço para acessar, organizar, gerenciar, tratar e disseminar a informação, conhecimentos e saberes*”. (BLATTMANN; SILVA, 2007).

⁷ *Web 3.0* se constitui como uma denominação para um período de evolução da Web marcado pela criação de ambientes informacionais altamente especializados e que só funcionarão efetivamente a partir da implantação da estrutura da *Web Semântica*. (SANTOS; ALVES, 2009).

1. Identificar, caracterizar e analisar o *MarcOnt Initiative* na sua essência, no âmbito das bibliotecas digitais e na *Web* na atualidade;
2. Caracterizar os padrões de representação e de descrição de recursos bibliográficos da área da Ciência da Informação, em especial o MARC 21;
3. Relacionar e analisar as similaridades e as diferenças do *MarcOnt* com o MARC 21, AACR2 e RDA;
4. Apontar as estruturas ontológicas da Catalogação Descritiva em códigos e padrões de metadados;
5. Propor uma arquitetura para a descrição bibliográfica semântica, que garanta a interoperabilidade em ambientes informacionais digitais.

1.4 Metodologia



No desenvolvimento de um arcabouço teórico e metodológico sobre as ferramentas tecnológicas disponíveis para a construção de formas de representação de recursos informacionais, no âmbito digital, o presente trabalho caracteriza-se por ser uma pesquisa de análise exploratória e descritiva do tema (CERVO; BERVIAN, 2003), na tentativa de identificar o embasamento conceitual subjacente às ferramentas para a padronização da descrição dos aspectos de forma e do conteúdo dos recursos bibliográficos, no âmbito digital, na abordagem da Ciência da Informação.

Assim, o referencial teórico publicado sobre o tema, oferece elementos para fundamentar e embasar a pesquisa visando-se ao alcance do principal objetivo estabelecido neste trabalho. A adoção dessa metodologia permitiu abordar os aspectos envolvidos nos objetivos específicos definidos anteriormente e a proporcionar a compreensão e a concretização do objetivo geral proposto. O método de coleta de dados escolhido é a pesquisa bibliográfica e documental.

Para a fundamentação metodológica são considerados:

a-) A proposta perspectivista de Peterson (1996) para a análise das formas de representação adotadas para a recuperação de documentos em acervos específicos, em uma

analogia entre a estrutura das linguagens computacionais e a estrutura do pensamento, no processo de resolução de problemas;

b-) A lógica de descrição, proposta pelos instrumentos de representação descritiva documentária, da área de Biblioteconomia. Como orientações visando-se à construção dos elementos de um processo de utilização e reutilização das informações, com o objetivo de, a partir da manipulação de formas de representação bibliográficas e da interação de sujeitos com o ambiente informacional, por meio da representação formal de documentos que garante a interoperabilidade entre os sistemas informacionais.

1.4.1 Procedimentos metodológicos

Os procedimentos metodológicos para a efetivação dessa investigação podem ser assim delineados:

- Levantamento bibliográfico em níveis nacional e internacional em fontes bibliográficas, conforme apresentado no Quadro 1:

Quadro 1: Pesquisa em fontes primárias e secundárias.

Primárias			
Livros	Periódicos	Anais de congressos	Teses e documentos eletrônicos da Internet, entre outros documentos congêneres;

Secundárias				
Bases de dados textuais e referenciais, como:				
Library and Information Science Abstracts (LISA) <ul style="list-style-type: none"> • Abrangência temática: cobertura internacional na área de Ciência da Informação. • Vinculação institucional: UNESP e Portal de Periódicos CAPES. • Tipos de documentos indexados: índices e resumos de periódicos da Ciência da Informação. • Período abrangido da base: 1969 até a presente data. 	Information Science & Technology Abstracts (ISTA) <ul style="list-style-type: none"> • Abrangência temática: Ciência da Informação. • Vinculação institucional: UNESP e Portal de Periódicos CAPES. • Tipos de documentos indexados: índices e resumos de artigos de revistas especializadas, além de livros, relatórios de pesquisa e análises de conferências e patentes. • Período abrangido da base: meados da década de 1960 	Library, Information Science & Technology Abstracts with Full Text (LISTA) <ul style="list-style-type: none"> • Abrangência temática: Ciência da Informação, Biblioteconomia, Classificação, Catalogação, Bibliometria, Recuperação de informações on-line, Gestão de informações. • Vinculação institucional: UNESP e Portal de Periódicos CAPES. • Tipos de documentos indexados: revistas científicas seletivas, além de livros, relatórios de pesquisas e protocolos. • Período abrangido da base: meados da década de 1960. 	Library Literature and Information Science Full Text <ul style="list-style-type: none"> • Abrangência temática: Biblioteconomia, Ciência da Informação e Editoração. • Vinculação institucional: UNESP e Portal de Periódicos CAPES. • Tipos de documentos indexados: periódicos, resenhas de livros, livros, teses, trabalhos de congressos e outros materiais. • Período abrangido da base: 1984 até o presente. 	Wilson Library Literature and Information Science Full Text <ul style="list-style-type: none"> • Abrangência temática: Ciência da Informação, Bibliotecas, Automação, Catalogação, Classificação e Pesquisa eletrônica. • Vinculação institucional: UNESP. • Tipos de documentos indexados: resumos e texto completo de periódicos, teses e livros. • Período abrangido da base: 1984 até o presente

Fonte: Elaborado pelo autor.

As estratégias de busca para cada base de dados foram realizadas na opção busca avançada (*Advanced Search*), pois esta possibilita a busca em campos específicos da base, tais como, palavras-chave (*Keywords*) e descritores (*Descriptors*) do *thesaurus* e com o operador *booleano AND*, a fim de obter a especificidade da busca pretendida, nos idiomas português, inglês e espanhol, a partir dos termos: *MarcOnt AND MARC*; *Catalogação AND MARC*; *Catalog*⁸ AND MARC*; *Catalogación AND MARC*; *FRBR AND Catalogação*; *FRBR AND Catalog**; *FRBR AND Catalogación*; *RDA AND Catalogação*; *RDA AND Catalog**; *RDA AND Catalogación*; *MarcOnt AND Catalogação*; *MarcOnt AND Catalog**; *MarcOnt AND Catalogación*; *MarcOnt AND Web Semântica*; *MarcOnt AND Semantic Web*; *MarcOnt AND Web Semântica*; *MARC AND FRBR*; *MARC AND RDA*; *MARC AND Metadados AND Catalogação*; *MARC AND Metadata AND Catalog**; *MARC AND Metadatos AND Catalogación*; *Catalogação AND Web 2.0*; *Catalog* AND Web 2.0*; *Catalogación AND Web2.0*; *Metadados AND Web Semântica AND Web 2.0*; *Metadata AND Semantic Web AND Web 2.0*; *Metadatos AND Web Semântica AND Web 2.0*. *Biblioteca* Digita* AND*

⁸ Adotou-se como estratégia o truncamento de alguns termos da busca como: *Catalog** no idioma inglês, pois no *thesaurus* das bases aparece a variação *Cataloguing* e *Cataloging*; o mesmo critério *Biblioteca* Digit** para *Biblioteca Digital* e *Bibliotecas Digitais*, em português e espanhol, e ainda *Digital Librar** para *Digital Library* e *Digital Libraries*, no idioma inglês. O termo *ontologia* também foi utilizado como *ontolog** para os idiomas em inglês e espanhol.

Metadados; Digital librar* AND Metadata; Biblioteca* Digita* AND Metadatos; Biblioteca* Digita* AND MARC; Digital librar* AND MARC; Biblioteca* Digita* AND MARC; Biblioteca* Digita* AND Web Semântica AND Metadados; Digital Librar* AND Semantic Web AND Metadata; Biblioteca* Digita* AND Web Semántica AND Metadatos; Metadata AND Ontolog* AND Semantic Web; Catalog* AND Metadata AND Ontolog*; RDF AND Ontolog* AND Semantic Web; RDF AND Metadata AND Ontolog*; FRBR AND Catalog* AND Ontolog*; Metadata AND FRBR AND Ontolog*.

A forma de busca e obtenção da literatura científica se deu por meio do acesso ao *site* do Conselho de Reitores das Universidades Estaduais Paulistas (CRUESP) que integra o Sistema de Bibliotecas da UNESP, USP e UNICAMP, uma vez que este mantém o intercâmbio institucional para empréstimo de materiais bibliográficos. Destaca-se também o Serviço de Comutação Bibliográfico (COMUT) disponível na biblioteca da unidade (UNESP/Marília), as Bibliotecas Digitais de Teses e Dissertações, Periódicos CAPES, UnibibliWeb (Portal de *e-books*), Scielo (periódicos científicos nacionais e internacionais) que podem ser acessados por meio da rede digital (Internet).

- A seleção dos documentos obedeceu aos critérios de pertinência com relação aos objetivos específicos desta pesquisa, tendo como base, a partir da análise entre a ontologia MarcOnt e o padrão de metadados MARC 21, a necessidade de verificação de elementos de interoperabilidade no processo de catalogação, que potencializa os processos de recuperação da informação em ambientes informacionais digital. Para tanto, foram delimitados os documentos referentes aos idiomas português, inglês e espanhol, com período de publicação limitado aos últimos dez anos, apenas como abordagem inicial, não havendo limitação cronológica para referências citadas nos documentos selecionados. Justificam-se os últimos dez anos e os idiomas mencionados, por se tratar de um período marcado por estudos realizados pela *International Federation of Library Associations and Institutions* (IFLA), sobre a necessidade de um reexame dos registros bibliográficos de modo a atender a uniformidade da descrição bibliográfica, e, por se considerar a década de grandes transformações na Catalogação, o que resultou numa acentuada produção científica nos idiomas supracitados.
- Leituras e documentação dos textos selecionados, que possibilitaram a criação de um referencial teórico com o qual foi possível obter subsídios para um maior entendimento e

compreensão mais detalhados do MarcOnt e do MARC 21 e as metodologias para a descrição de recursos informacionais, no âmbito das bibliotecas digitais e da *Web*, na abordagem da Ciência da Informação, e o impacto e as tendências da prática da Catalogação no século XXI.

1.4.2 Forma de análise dos resultados

Para que os objetivos propostos pela pesquisa pudessem ser alcançados, e para atender aos critérios de sistematicidade, consistência e replicação de um trabalho científico, os resultados obtidos estão sistematizados sob os focos:

1. Identificação e caracterização do MarcOnt Initiative no âmbito da Ciência da Informação, a fim de contextualizar o objeto de estudo da pesquisa e sua relevância e contribuição científica para a área.

A partir da literatura científica disponível, procura-se identificar os conceitos e as características subjacentes à ontologia MarcOnt, a fim de construir o *corpus* teórico da pesquisa e a sistematização das características fundamentais extraídas da literatura e suas interfaces com o padrão de metadados MARC 21.

2. Sistematização do MarcOnt e do formato de intercâmbio MARC21.

A partir da caracterização do MarcOnt *Initiative*, faz-se o estudo das similaridades e das diferenças desse com o padrão de metadados MARC 21 componente da Biblioteconomia e da Ciência da Informação, a partir dos esquemas de descrição bibliográficos, como o AACR2 e RDA, comparando-os conceitualmente com os principais recursos informacionais encontrados na ontologia MarcOnt.

3. Proposta de uma arquitetura para a descrição bibliográfica dos aspectos semânticos de forma e conteúdo dos recursos informacionais na perspectiva da interoperabilidade entre ambientes informacionais digitais na atualidade.

Com base na análise sistematizada da fundamentação teórica e dos instrumentos de descrição da Ciência da Informação, em consonância com as novas transformações ocorridas na Catalogação Descritiva no século XXI, comprovou-se que a ontologia MarcOnt não se adequava aos propósitos determinados pela Catalogação Descritiva, nas formas de representação e de descrição padronizada de recursos bibliográficos, fazendo-se necessário o estudo e a identificação de elementos de interoperabilidade no processo de catalogação, que permite a interoperabilidade em ambientes informacionais digitais. Assim, buscou-se na literatura científica, no domínio bibliográfico, sobretudo, com o desenvolvimento dos FRBR e o novo código de catalogação (RDA), quais os requisitos que os ambientes informacionais digitais atualmente devem atender, sobretudo na estruturação (camadas de representação e descrição) dos dados bibliográficos, potencialmente utilizados para formatos na *Web*, no estabelecimento efetivo da interoperabilidade.

Dessa maneira, apresenta-se a proposta de uma arquitetura com diretrizes funcionais, a fim de identificar teórica e metodologicamente os relacionamentos entre elementos de metadados, naquilo que envolve a descrição bibliográfica e a interoperabilidade semântica entre padrões de metadados heterogêneos e sua interpretação pelas máquinas e ainda, a percepção dos usuários sobre o recurso informacional no ambiente informacional digital.

1.5 Justificativa

Um dos aspectos que motivaram e nortearam o desenvolvimento desta tese foi a familiaridade do autor desde a pesquisa na graduação em Biblioteconomia no ano de 2002, no tema Catalogação Automatizada e formato de intercâmbio MARC 21, o que impulsionou um estudo mais aprofundado sobre a temática aqui tratada.

Em 2005, no curso de especialização *Uso Estratégico das Tecnologias em Informação*, oferecido pela Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP/Marília, o autor se deparou com questões relacionadas às novas abordagens da representação e da descrição de recursos informacionais em ambientes informacionais digitais, tais como a *Web Semântica* e as Bibliotecas Digitais e os novos rumos da Catalogação aplicada a essas ambiências.

Em 2006, no curso de Mestrado em Ciência da Informação, da mesma instituição, foi possível identificar dentre algumas perspectivas tecnológicas, o surgimento de novas tendências para a prática da descrição bibliográfica, denominado *MarcOnt Initiative*, bem como a necessidade de estudá-lo de forma mais aprofundada, destacando sua relevância para as questões de representação e de descrição de recursos informacionais e a integração de esquemas de metadados heterogêneos.

A iniciativa *MarcOnt* nessa tese foi o objeto de estudo que permitiu que se comprovasse a necessidade de elementos de interoperabilidade no processo de catalogação. *MarcOnt* foi eleito porque propagava ter uma estrutura que garantiria a interoperabilidade entre recursos por meio de um “conversor” independente da estrutura original de descrição do recurso.

Partindo do princípio de que cada vez mais informações são produzidas, principalmente no meio digital, tais ambientes passam por uma transição de readequação dos seus métodos que possibilitem a descrição, a representação e a organização de seus conteúdos, de modo mais significativo, eficiente e preciso.

Tal cenário requer uma preocupação com a identificação e o desenvolvimento de normas e padrões para a representação das informações que facilitem sua descrição e localização, e ainda que proporcione a interoperabilidade entre os sistemas informacionais.

Ainda que amplamente discutido em nível internacional, por exemplo, na Biblioteca Digital Semântica *JeromeDL*, e contando com o apoio financeiro de renomados órgãos de fomento, no Brasil não existem relatos de discussão a respeito do *MarcOnt Initiative*, no processo de recuperação da informação e identificação do recurso informacional em ambientes digitais.

A atualidade do assunto e sua repercussão na comunidade científica, atrelado ao desejo e a necessidade de uma investigação científica de forma mais aprofundada no tema MARC e formas de representação no âmbito das bibliotecas digitais na era da *Web Semântica*, também é levado em consideração, uma vez que a maior base bibliográfica do Brasil (Catálogo Coletivo Bibliodata) – está cunhada neste formato e internacionalmente, a *Library*

of Congress – como um grande repositório de registros bibliográficos e catalográficos, também em formato MARC 21.

Pensando nisso, acredita-se que esta seja uma pesquisa de relevância por contribuir para o desenvolvimento de referencial teórico e metodológico na disciplina de Catalogação da área de Biblioteconomia, uma vez que é possível identificar uma escassez de literatura sobre a temática de representação descritiva, devendo tal assunto ser repensado, principalmente com a criação dos Requisitos Funcionais para Registros Bibliográficos (FRBR) e o novo código de catalogação denominado *Resource Description and Access* (RDA), considerados marcos fundamentais na perspectiva da descrição bibliográfica.

Verifica-se, preliminarmente, a falta de publicações do assunto em língua portuguesa, destacando dessa forma, a relevância da pesquisa por contribuir como uma fonte de informação bibliográfica; além disso, a relevância social na realização desta pesquisa está em fornecer subsídios teóricos e metodológicos, para que profissionais da área de Ciência da Informação possam desenvolver métodos e técnicas mais eficazes de representação para a localização, a recuperação, o uso e o (re)uso da informação e que ainda permita a interoperabilidade semântica em ambientes informacionais digitais.

Acredita-se, que este estudo possibilite ir ao encontro dos principais atores envolvidos nesse cenário, em especial, os usuários, sujeitos psicossociais que se apropriarão e farão uso dos recursos informacionais disponíveis nos variados ambientes informacionais digitais. Destaca-se, que os elementos que compõem as apresentações das representações bibliográficas em ambientes informacionais digitais, são organizados com objetivos distintos para a motivação do interesse do usuário para acesso, uso e (re)uso das informações.

Além disso, justifica-se o desenvolvimento da pesquisa, no que tange em acompanhar os avanços da área de Ciência da Informação sobre a temática abordada e a questão da padronização da representação e da descrição dos recursos informacionais e seus conteúdos que garantirão a interoperabilidade semântica entre ambientes e sistemas informacionais digitais e, ainda, como uma base consolidada para a descrição bibliográfica nacional e internacionalmente.

Atenta-se também, como uma relevante investigação de caráter teórico e metodológico sobre a descrição bibliográfica e o tratamento descritivo da informação atualmente, por sua contribuição à linha de pesquisa a qual está inserida Informação e Tecnologia do Programa de Pós-Graduação em Ciência da Informação – UNESP/Marília.

1.6 Estrutura da Tese

Para uma melhor compreensão e organização desta tese, a presente investigação, além do presente capítulo (1) que trata da **Introdução** e que aborda questões iniciais, como o contexto que esta se insere, o problema da pesquisa, a hipótese, a proposição, a tese, os objetivos (geral e específico), a metodologia e os procedimentos metodológicos, as formas de análise dos resultados, a justificativa, a relevância social e a relevância para a área de Ciência da Informação, está estruturada da seguinte forma:

CAPÍTULO 2 *ONTOLOGIAS E METADADOS: aproximações sinérgicas para o estabelecimento da interoperabilidade em ambientes digitais*

Este capítulo aborda em que medida as ontologias e os metadados podem possibilitar novas perspectivas para as transformações na descrição bibliográfica, nos modelos de relacionamentos entre esquemas de metadados heterogêneos e o seu encaminhamento para o desenvolvimento de ambientes informacionais digitais melhor estruturados.

CAPÍTULO 3 *MARCONT: um diálogo possível com MARC 21?*

Trata-se do estudo sistematizado e minucioso do MarcOnt em sua essência, as classes e as propriedades encontradas nos recursos informacionais, bem como as principais tecnologias para seu desenvolvimento. Destaca-se ainda, o estudo realizado com os instrumentos de descrição bibliográfica da Biblioteconomia e da Ciência da Informação, o *Anglo-American Cataloging Rules* (Código de Catalogação Anglo- Americano) - AACR2 - e o *Resource Description Framework* (RDA), e também o padrão de metadados MARC 21,

num comparativo conceitual entre os principais termos (elementos) da ontologia MarcOnt. Aponta-se ainda que a ontologia MarcOnt e a arquitetura de metadados RDF são indicadas para o estabelecimento efetivo da interoperabilidade semântica entre os sistemas e ambientes informacionais digitais.

CAPÍTULO 4 RESOURCE DESCRIPTION FRAMEWORK (RDF): promessa para interoperabilidade em ambientes informacionais digitais

Após a primeira etapa da pesquisa de construção de seu *corpus* teórico e das análises dos instrumentos de descrição e representação de recursos informacionais, o capítulo 4 aborda como a arquitetura de metadados RDF poderá proporcionar, juntamente com os modelos conceituais, sobretudo, no domínio bibliográfico, a infraestrutura dos novos ambientes informacionais digitais atrelado aos metadados no estabelecimento da interoperabilidade, atuando como o núcleo de armazenamento de descrições bibliográficas semânticas.

CAPÍTULO 5 CATALOGAÇÃO DESCRITIVA NO SÉCULO XXI: proposta de um modelo para a descrição bibliográfica semântica em ambientes informacionais digitais interoperáveis

Com os avanços tecnológicos no contexto da Catalogação Descritiva, procura-se verificar as relações e as correspondências de forma explícita entre as ontologias e os esquemas de metadados, e a proposta de uma arquitetura conceitual para a construção de descrição bibliográfica semântica, fundamentada nos princípios da representação descritiva e suas imbricações nos atuais ambientes informacionais digitais.

CAPÍTULO 6 CONSIDERAÇÕES FINAIS

Esse capítulo apresenta as observações e as reflexões acerca das tendências na área da Catalogação Descritiva e seu impacto na construção de novos ambientes informacionais digitais estruturados e interoperáveis, na perspectiva da Ciência da Informação.

REFERÊNCIAS - São apresentadas as referências que constituem e formam o arcabouço teórico e metodológico dessa pesquisa.



CAPÍTULO 2



ONTOLOGIAS E METADADOS: APROXIMAÇÕES SINÉRGICAS PARA O ESTABELECIMENTO DA INTEROPERABILIDADE EM AMBIENTES DIGITAIS

2 ONTOLOGIAS E METADADOS: aproximações sinérgicas para o estabelecimento da interoperabilidade em ambientes digitais

A condição fundamental das possibilidades de um saber pertinente é o saber dos pressupostos fundamentais de todo saber. *Heidegger*

Para se contextualizar a questão de pesquisa, faz-se necessário destacar alguns marcos que ajudarão a compreender as bases conceituais, as relações e as aproximações que as áreas de metadados e de ontologias possuem, bem como a sua sincronicidade que permite/conduz à interoperabilidade semântica entre ambientes informacionais digitais.

2.1 Ontologias: algumas considerações

Heterogeneidade é uma das principais características das coleções do patrimônio cultural. A diversidade documental pode compreender textos, escritos em diferentes materiais (suportes de armazenamento), pinturas, fotografias, objetos 3D, gravações de som, mapas e até mesmo objetos digitais. Além disso, os objetos estão atrelados aos acontecimentos sociais e históricos que têm lugar ao longo dos tempos. Outro fato que se deve levar em consideração é a estrutura (sintaxe), semânticas diversas e os vários tipos de relações entre os objetos dessas coleções.

Assim, tornar acessíveis os recursos do patrimônio cultural exige ou compreende a adoção de esquemas e/ou padrões de metadados ricos⁹ na semântica e estruturas, que sejam capazes de cobrir a heterogeneidade material e a variedade da memória dos ambientes informacionais e/ou instituições (bibliotecas, arquivos, museus etc.).

⁹ Os padrões de metadados ricos apresentam um esquema de metadados complexo que reflete o grau de especificidade descritiva de um domínio específico. Ex.: formato/padrão de metadados MARC 21.

Pensando na variedade de esquemas de metadados do patrimônio cultural, que são muitas vezes semanticamente relacionados uns aos outros, e no aumento da demanda de usuários para o acesso global a coleções altamente distribuídas, heterogêneas e dinâmicas, a ênfase é dada nas questões de integração e interoperabilidade de recursos. A interoperabilidade é necessária não só sintática e em nível do sistema, mas também ao nível mais complexo semântico.

No contexto da integração de dados, a heterogeneidade é fator essencial a ser tratado, tendo em vista que as fontes a serem integradas foram desenvolvidas em contextos distintos, por pessoas e organizações diferentes. Assim, são encontrados modelos de dados diversos, representações diferentes para os dados, terminologias conflitantes etc.

A integração de dados, principalmente em ambientes informacionais digitais, tem sido objeto de investigação e pesquisa, dinâmica e desafiadora durante muitos anos, com o intuito de fornecer aos usuários interfaces uniformes e mais amigáveis de acesso, relacionando e combinando os diferentes recursos.

As fontes de dados podem ser heterogêneas na sintaxe, no esquema ou na semântica, tornando uma tarefa difícil à interoperabilidade dos dados. A heterogeneidade sintática é causada pelo uso de diferentes modelos ou linguagens. A heterogeneidade esquemática é resultado das diferenças estruturais. Enquanto isso, a heterogeneidade semântica é causada pelos diferentes significados ou interpretação dos dados em vários contextos. Para alcançar a interoperabilidade dos dados é necessária a eliminação de sua heterogeneidade. (CRUZ; XIÃO, 2005, tradução nossa).

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.

Faz-se uma ressalva que a iniciativa MarcOnt vai ao encontro, ou poderia ser considerada uma tentativa de interoperabilidade semântica entre sistemas e ambientes

informacionais e esquemas ou padrões de metadados heterogêneos, tendo como *core* a linguagem RDF (*Resource Description Framework*) para a compilação dos dados e o armazenamento de descrições semânticas, conforme pode ser vista na arquitetura de mediação de serviços MarcOnt, apresentada no capítulo 3.

Recentemente, o interesse em pesquisas em muitas comunidades e disciplinas está passando por uma fase de transição, ou seja, percebe-se uma mudança da integração de dados para a **integração semântica de dados**, como por exemplo, em aplicações de sistemas de informações geográficas, gerenciamento de informações médicas, integração de informações empresariais e *e-Gov*¹⁰.

A integração semântica de dados seria “o processo de utilização de uma representação conceitual dos dados e seus relacionamentos, com o objetivo de eliminar as possíveis heterogeneidades de compreensão e interpretação”. (CRUZ; XIÃO, 2005, p. 2, tradução nossa). Ela ainda pode ser entendida como aquela que inclui a semântica das requisições de informação do usuário e das fontes de informação.

Nessa perspectiva Cruz e Xião (2005) apontam que a heterogeneidade semântica pode ser classificada em:

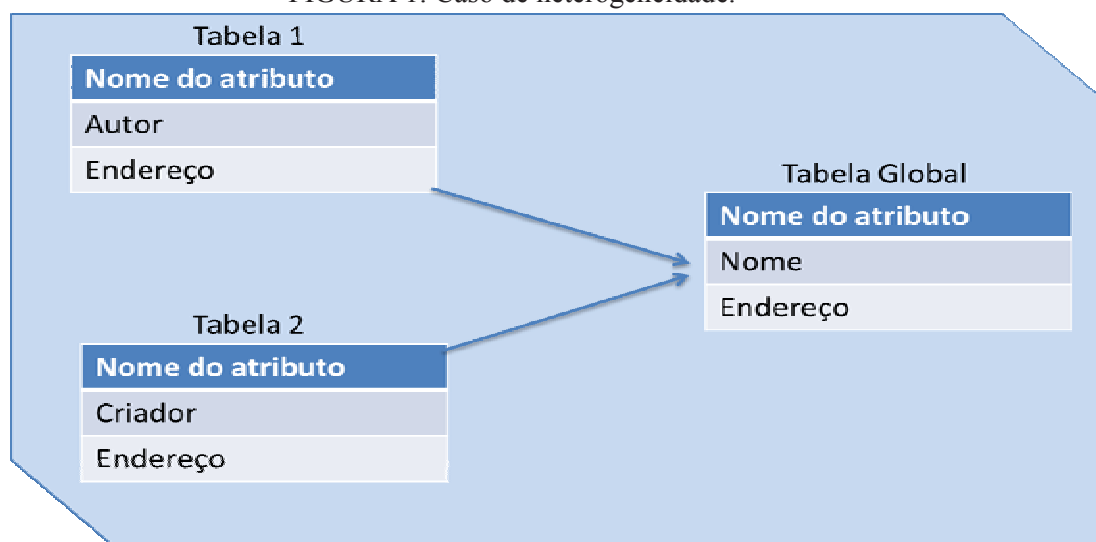
- Conceitos semanticamente equivalentes: termos diferentes são usados para referenciar o mesmo conceito; propriedades diferentes são modeladas de forma diferente por sistemas distintos;
- Conceitos semanticamente não-relacionados: o mesmo termo usado em sistemas distintos para denotar coisas completamente diferentes;

¹⁰ O E-GOV pode ser entendido como uma das principais formas de modernização do estado e está fortemente apoiado no uso das novas tecnologias para a prestação de serviços públicos, mudando a maneira com que o governo interage com os cidadãos, empresas e outros governos. O conceito não se restringe a simples automação dos processos e disponibilização de serviços públicos através de serviços on-line na Internet (ABRAMSON; MEANS, 2001), mas sim na transformação da maneira com que o governo, através da TIC, atinge os seus objetivos para o cumprimento do papel do estado. O governo eletrônico engloba, principalmente, três tipos de transações, a saber: G2G (forma reduzida da expressão em inglês “*Government to Government*”) trata-se de uma relação intra ou inter-governos (Gestão interna: Sistemas Internos - Intranet); G2B (forma reduzida da expressão em inglês “*Government to Business*”) caracteriza-se por transações entre governos e fornecedores (integração com a cadeia de valor: Sistemas Estruturadores - Extranet); G2C (forma reduzida da expressão em inglês “*Government to Citizen*”) envolvendo relações entre governos e cidadãos (interação com cidadão: Portal de atendimento - Internet). (FERNANDES, 2000).

- Conceitos semanticamente relacionados: generalização e especificação; conceitos sobrepostos; conceitualizações diferentes.

A figura 1 apresenta um caso onde poderia se pensar na adoção de uma semântica maior para auxiliar a descrição de um recurso. Na Tabela local 1, o atributo nome representa o nome de um *autor*; entretanto, a Tabela local 2, o mesmo nome é representado pelo atributo *criador*. Um maior nível semântico pode ser agregado a essas tabelas locais, podendo indicar que ambos os atributos (tabelas 1 e 2) referem-se a nome, facilitando a integração dos mesmos em um atributo nome num esquema global.

FIGURA 1: Caso de heterogeneidade.



Fonte: Elaborado pelo autor.

Chateaubriand (1998) afirma que uma das tendências do século XXI é fazer a ontologia, derivar da semântica, uma vez que há uma relação entre a linguagem e a realidade. Como forma de possibilitar a adição de semântica aos metadados e a integração de esquemas, a literatura científica aponta e tem sido proposto o uso de ontologias, em decorrência do seu potencial para descrever a semântica de fontes de informação distintas e como forma de resolver os problemas de heterogeneidade na *Web*.

O advento da *eXtensible Markup Language* (XML¹¹) criou uma plataforma sintática para a padronização e o intercâmbio de dados na *Web*. No entanto, a heterogeneidade

¹¹ XML: uma metalinguagem e um simples dialeto de *Standard Generalized Markup Language* (SGML) desenvolvida pelo W3C para o processamento de documentos contendo informações estruturadas.

esquemática dos dados pode persistir, dependendo do *schema* XML¹² utilizado, por exemplo, níveis hierárquicos. Um ponto importante que Cruz e Xião (2005) destacam é que a heterogeneidade semântica pode persistir, mesmo se a sintaxe e os esquemas não ocorrerem, por exemplo, nomear os conceitos de maneira diferente.

No centro da integração semântica dos dados está o conceito de ontologia, ou seja, uma especificação explícita de uma conceituação compartilhada, com o propósito de facilitar o compartilhamento do conhecimento em áreas distintas para o (re)uso dos dados.

Uma ontologia define os termos usados para descrever e representar uma área do conhecimento. Ontologias são usadas por pessoas, bancos de dados e aplicações que necessitam compartilhar informações de domínio (um domínio é apenas um assunto específico de uma área ou área de conhecimento, como a medicina, gestão financeira etc.). Ontologias incluem definições de conceitos básicos processáveis pelo computador no domínio e os relacionamentos entre eles. As ontologias codificam o conhecimento em um domínio e também o conhecimento que cobrem os domínios. Dessa forma, faz com que o conhecimento seja reutilizável. (OWL, 2004, tradução nossa).

Assim, “ontologias são usadas para capturar o conhecimento do domínio da aplicação e usá-lo para responder aos requisitos de informação da melhor maneira possível” (CRUZ; XIÃO, 2005, p. 2, tradução nossa).

Uma das principais infraestruturas da *Web* Semântica, em relação às questões da integração semântica são as ontologias que se apresentam como instrumentos para a descrição de recursos informacionais. Segundo Cunha (2002), uma ontologia fornece um entendimento comum e compartilhado de um domínio, que pode ser comunicado através de pessoas e sistemas de aplicação, tornando-se elemento-chave para o desenvolvimento da *Web* Semântica.

¹² *XML schema* pode ser considerado como um esquema XML, ou seja, é um documento que define, em um método formal, a estrutura, o conteúdo e as semânticas de documentos XML.

A ontologia tem um papel crucial no sentido que permite o acesso, a interoperação e a comunicação baseados em conteúdo, fornecendo à *Web* um nível de serviço qualitativamente novo, que consideramos na *Web Semântica*. Ela une em rede incríveis porções do conhecimento humano, complementando-as com capacidade de processamento de máquina. (CUNHA, 2002, p. 14).

As ontologias fornecem um entendimento compartilhado de um domínio de interesse para a comunicação entre humanos e agentes de *softwares*, normalmente representado em uma linguagem processável pela máquina.

Nesse sentido, as ontologias, com base na literatura científica do campo da Ciência da Informação, oferecem soluções para o problema da heterogeneidade semântica, podendo ser usadas na integração de arquiteturas de metadados, como um esquema global para o mapeamento de diferentes fontes de metadados. Em comparação a outros esquemas, as ontologias conceituam domínios de interesses particulares e expressam rica semântica. (CRUZ; XIÃO, 2005).

De acordo com Corcho, Fernández-Lopez e Gómez-Pérez (2001, tradução nossa), ontologias são constituídas de 5 (cinco) componentes centrais, dentre os quais se destacam: classes, relacionamentos, funções, axiomas e instâncias e podem ser entendidas no domínio da Catalogação Descritiva, de acordo com o Quadro 2.

Quadro 2: Analogia da ontologia com a Catalogação Descritiva.

Classes	são comumente organizadas em taxonomias. As classes ou conceitos são utilizados num sentido mais geral, ou seja, podendo ser o conceito propriamente dito sobre alguma coisa ou a descrição de tarefas, funções, ações, estratégias, processo de raciocínio etc. Um exemplo na Catalogação Descritiva pode ser representado pelo conceito da unidade informacional (livro, periódico, mapa etc.).
Relações	representam o tipo de envolvimento ou interação entre os conceitos de um domínio. Elas podem ser formalmente definidas entre subconjuntos, tal como, as relações do tipo <u>subclasse de</u> ou <u>conectada a</u> . No domínio bibliográfico podem-se ver as relações entre os elementos descritivos de um padrão de metadados, por exemplo, o MARC 21.
Funções	apresentam-se como um caso especial de relação onde o enésimo elemento da relação é único para os n-1 elementos precedentes. Exemplo: antecedente e causa.
Axiomas	são utilizadas para modelar as sentenças que são sempre verdadeiras e aplicadas de acordo com várias propostas: definir exatamente o significado dos componentes de uma ontologia, as restrições complexas nos valores dos atributos etc.; de forma a verificar a correção das informações especificadas na ontologia, ou ainda, deduzindo novas informações. Exemplo: as regras e/ou esquemas de descrição bibliográfica, AACR2.
Instâncias	são empregadas para representar elementos específicos, ou seja, os dados propriamente ditos. Na representação bibliográfica têm-se os dados de descrição, tais como, autor, título, ano de publicação etc.

Fonte: Adaptado de Corcho, Fernández-Lopez e Gómez-Pérez (2001).

No escopo das investigações sobre ontologias, no que tange aos tipos de ontologias, a literatura científica apresenta algumas classificações, conforme pode ser visualizado no Quadro 3.

Quadro 3: Classificação de ontologias.

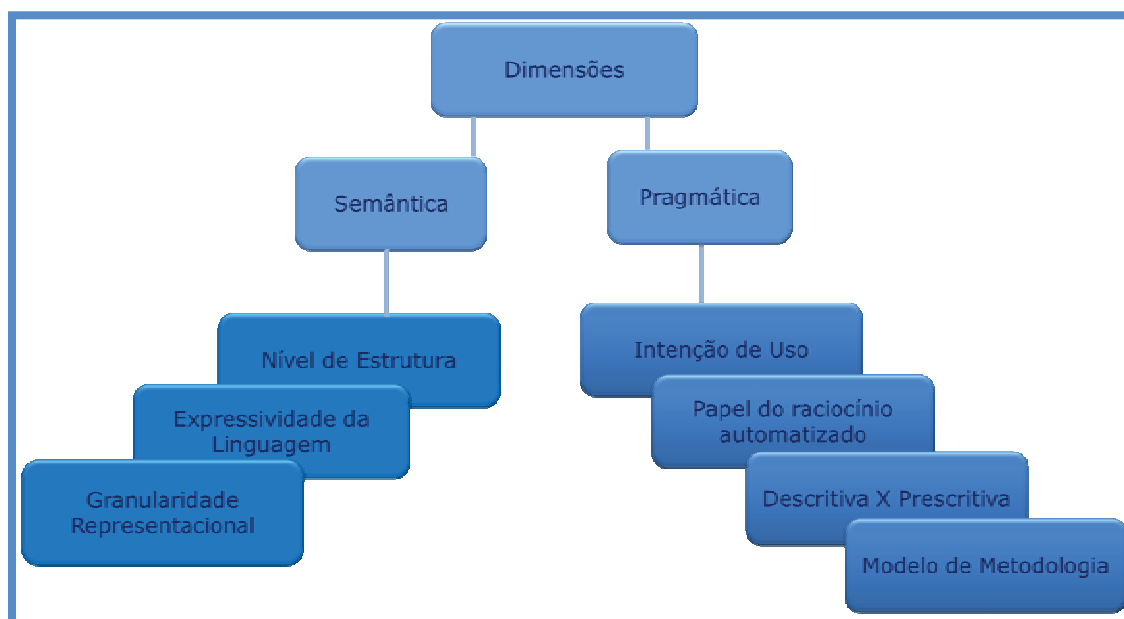
Ano	Autor	Classificação de ontologias
1997	VAN HEIJST; SCHREIBER; WIELINGA	Dimensão quanto à estrutura e assunto da conceitualização
1999	USCHOLD; JASPER	Dimensão quanto ao cenário para aplicação
2001	HAAV; LUBI	Dimensão quanto ao tipo de classes presentes
2003/2006	ALMEIDA; ALMEIDA; BAX	Dimensão quanto à conceitualização
2004	ALBUQUERQUE; KERN	Dimensão quanto ao conteúdo
2004	ALBUQUERQUE; KERN	Dimensão quanto à profundidade ontológica
2006	ALMEIDA	Dimensão quanto ao grau de formalidade
2007	GRUBER et al.	Dimensões semântica e pragmática

Fonte: elaborado pelo autor.

Dentre as tipologias supracitadas, Gruber et al. (2007) propõem um *framework* para a área de ontologias estabelecendo uma base conceitual no campo da Ciência da Informação e da Ciência da Computação, identificando as convergências e as divergências entre as ontologias.

No âmbito da dimensão das ontologias, Gruber et al. (2007) apontam duas abordagens, a semântica e a pragmática, conforme pode ser observado na figura 2:

FIGURA 2: Abordagens e tipologias de ontologias segundo Gruber et al. (2007).



Fonte: Adaptado de Gruber et al. (2007, tradução nossa).

Gruber et al. (2007, tradução nossa) apresentam na dimensão semântica três níveis os quais estão classificadas as ontologias.

- Nível de Estrutura: aparecem semelhantes aos conceitos de dados estruturados¹³ e não estruturados¹⁴ no campo da Ciência da Computação. As ontologias podem ser altamente estruturadas, semiestruturadas ou pouco estruturadas.

Uma ontologia que especifica conceitos definidos formalmente, tal como abstrações matemáticas é altamente estruturada, enquanto uma ontologia que especifica conceitos muito gerais, tais como, documentos e hiperligações, é pouco estruturada. Muitas ontologias são semiestruturadas, contendo uma mistura de definições formais e informais de conceitos e de relacionamentos. Por exemplo, uma ontologia bibliográfica para dados sobre livros pode conter o conceito de data com restrições formais sobre a noção de tempo previsto (altamente estruturada) e o conceito de título do livro que somente é identificado como uma sequência (*string*) de texto (pouco estruturada).

¹³ Os dados estruturados são aqueles armazenados num sistema de gerenciamento de banco de dados, com esquema rígido. Um exemplo no campo da biblioteca seria um catálogo bibliográfico estruturado no formato MARC 21.

¹⁴ Os dados não estruturados seriam os dados que não possuem nenhum tipo de estrutura determinada, por exemplo, uma imagem, um texto, um vídeo etc.

Para Gruber et al. (2007) uma outra maneira de pensar o nível da estrutura de uma ontologia é o grau que um vocabulário é restringido e pode suportar computação: dados estruturados, tais como banco de dados numérico relacional são altamente restringidos e suportam um alto grau de computação; entretanto, um banco de dados textual com hiperligações, em linguagem natural, possui poucas restrições que podem ser deduzidas ou inferidas a partir dos dados.

- Expressividade da Linguagem: Uma ontologia define seu vocabulário em alguma forma de representação. Ontologias diferem no poder expressivo da linguagem usada na especificação. Algumas conceitualizações requerem uma linguagem altamente expressiva para definir os conceitos, onde outros podem ser especificados com uma linguagem menos expressiva. A expressividade da linguagem tem relação com o nível da dimensão estrutural. Uma ontologia altamente estruturada e formal pode exigir uma linguagem com alta expressividade que comporte restrições lógicas ou matemáticas. Enquanto isso, uma ontologia pouco estruturada e informal pode ser expressa, a partir de uma simples lista de condições e definições numa linguagem natural, ou seja, uma linguagem com baixa expressividade. Por exemplo, indivíduos ou instâncias não podem ser expressos na lógica proposicional, enquanto o cálculo de predicados pode expressar as propriedades dos indivíduos.

- Granularidade¹⁵ Representacional: Enquanto a expressividade é uma característica da linguagem em que uma ontologia é dada, a granularidade é uma propriedade do conteúdo da própria ontologia, ou seja, representa o nível de detalhamento dos conceitos. A granularidade representacional especifica muito mais detalhes sobre as propriedades dos conceitos e como eles podem se relacionar uns com os outros.

¹⁵ Granularidade é “o nível de detalhamento o qual um objeto ou recurso informacional é visto ou descrito”. (BACA, 2008, p. 7, tradução nossa).

Na esfera da dimensão pragmática, Gruber et al. (2007) destacam quatro níveis para as ontologias:

- Intenção de Uso: as ontologias devem compartilhar bases de conhecimento¹⁶ para permitir a comunicação entre os agentes de *software*, para ajudar a integrar conjuntos de dados discrepantes, para representar um vocabulário em linguagem natural, para ajudar prover pesquisas baseadas em conhecimento, para fornecer um ponto de partida para a construção de sistemas de conhecimento, para fornecer uma estrutura conceitual para a indexação de conteúdos etc.
- Papel do raciocínio automatizado: o raciocínio automatizado pode variar do simples ao complexo.

No caso do simples, Gruber et al. (2007) afirmam que pode significar a interoperabilidade semântica, que requer que a linguagem com a qual o conteúdo é modelado seja baseada na lógica, ou que um motor especial de inferência/interpretação seja construído para conseguir interpretar um conteúdo. A primeira abordagem (a lógica) é baseada em princípios ou padrões.

A segunda abordagem (construção de um mecanismo específico para interpretação/inferência) é uma abordagem *ad hoc* e proprietária. No caso do raciocínio automatizado simples, uma máquina de busca pode ser capaz de fazer inferências, tais como as relações das subclasses (propriedades definidas na classe Pai podem ser herdadas pelas classes Filha), é a propriedade da transitividade.

O raciocínio automatizado mais complexo é geralmente expresso com regras dedutivas, isto é, regras de inferência ou expressões que combinam informações de toda ontologia.

¹⁶ O conceito de bases de conhecimento se disseminou graças ao desenvolvimento e à ampla utilização das TIC. Elas funcionam como facilitador à transferência do conhecimento explícito nas organizações. Davenport e Prusak (2001) enumeram os três tipos mais comuns de bases de conhecimento, denominadas de acordo com o tipo de conhecimento armazenado: bases de conhecimento externo (adquirido em fontes externas da organização); bases de conhecimento interno estruturado (relatório de pesquisa); bases de conhecimento interno informal (mais conhecidas como melhores práticas).

- Descritiva X Prescritiva: as descritivas são aquelas que frequentemente usam uma notação de caracterização mais livre (*looser*), permitindo objetos arbitrários no modelo, que não poderiam existir no mundo real, mas que são conceitos significantes para a comunidade de utilizadores. As prescritivas geralmente levam uma rígida noção na caracterização, declarando somente objetos que atualmente existem ou que representam tipos naturais ou tipos de coisas no mundo real e devem ser representados no conteúdo de um modelo de engenharia.

- Modelo de Metodologia: Nesse aspecto, os autores se referem se houve uma metodologia na construção da ontologia, que podem ser caracterizadas do tipo *bottom-up* ou *top-down*. *Bottom-up* é também denominada de empírica e tem forte ênfase em aspectos tais como, qualquer um que analise os dados resultantes da ontologia, compreenda sua semântica; ou capacitar pessoas voluntárias para caracterizar seus conteúdos como elas pessoalmente vêem, usando terminologia ou metadados e quaisquer relações estruturadas (ou não), que elas desejem usar, com o objetivo de que padrões de caracterização possam emergir ou ser preferidos por um grande grupo ou comunidade de pessoas.

Para essa tese utilizar-se-á a definição de ontologia proposta por Gruber (2009), pois suas imbricações nas áreas de Ciência da Informação e da Ciência da Computação corroboram a modelagem dos ambientes informacionais digitais e convergem para os aspectos de interoperabilidade.

Uma ontologia, para as ciências da computação e da informação, é uma especificação de uma conceitualização, que é o conjunto de idéias, conceitos, relacionamentos, ou outras abstrações que compõem um domínio de modelagem ou discurso. Uma ontologia define um vocabulário representacional para a conceitualização, e especifica restrições sobre o uso significativo deste vocabulário, de modo que os fatos sobre o domínio podem ser compartilhados, comunicados e fundamentados. (GRUBER, 2009, p. 3, tradução nossa).

Considera-se nessa pesquisa, a ontologia como um artefato tecnológico que propiciará a modelagem de ambientes informacionais digitais, através do estabelecimento conceitual na definição dos elementos de metadados, no processo da descrição bibliográfica, para uma compreensão/leitura da semântica para máquinas, oportunizando a interoperabilidade em ambientes informacionais digitais.

Vale destacar, nesse momento, que não é interesse dessa tese apontar profundamente os marcos conceituais e epistemológicos sobre ontologias, mas elucidar que elas podem propiciar a clareza nos relacionamentos entre sistemas e ambientes informacionais atreladas aos esquemas e/ou padrões de metadados, a fim de proporcionar a interoperabilidade, uma vez que tal estudo já fora realizado por diversos autores e em diferentes trabalhos.

Nesse sentido, o quadro 2 apresenta o arcabouço teórico e metodológico das investigações científicas contemporâneas com base na literatura nacional e internacional sobre ontologias e suas vertentes nas esferas da Ciência da Informação e da Ciência da Computação, destacando-as como requisitos para o estabelecimento e infraestrutura da *Web Semântica* e também para a interoperabilidade semântica, elencados de acordo com os autores: Alvarenga (2006); Feitosa (2006); Lima-Marques (2006); Campos, Campos e Campos (2006); Zhang, Cheng e Qu (2007); Pickler (2007); Castro (2008); Marcondes et al. (2008); Kim et al. (2008); Catarino e Baptista (2008); Kruk e McDaniel (2009); Kruk (2010); Ramalho (2010).

Quadro 4: Sistematização das investigações contemporâneas sobre ontologias sob o prisma da Ciência da Informação e Ciência da Computação.

Ano	Autor	Área	Estudo de ontologias com enfoque em:
2006	ALVARENGA, Lidia	Ciência da Informação	Organização e tratamento da Informação
2006	CAMPOS, Maria L. M.; CAMPOS, Maria L. de. A.; CAMPOS, Linair M.	Ciência da Informação	Representação do conhecimento
2006	FEITOSA, Ailton	Ciência da Informação	Representação da informação
2006	LIMA-MARQUES, Mamede	Ciência da Informação	Representação do conhecimento/Arquitetura da Informação
2007	PICKLER, Maria. E. V	Ciência da Informação	Representação da informação
2007	ZHANG, Xiang; CHENG, Gong; QU, Yuzhong	Ciência da Informação	Inteligência Artificial/ Métodos e Formalismos da Representação do Conhecimento
2008	CASTRO, Fabiano F. de.	Ciência da Informação	Representação da Informação
2008	CATARINO, Maria E.; BAPTISTA, Ana A.	Ciência da Informação/ Sistemas de Informação	Serviços e Sistemas de Informação para Web
2008	KIM, Hak L. et al.	Ciência da Computação	Representação do conhecimento
2008	MARCONDES, Carlos H. et al.	Ciência da Informação	Representação do conhecimento
2009	KRUK, Sebastian R.; McDANIEL, Bill	Ciência da Computação	Representação do conhecimento/Inteligência Artificial
2010	KRUK, Sebastian R	Ciência da Computação	Representação do conhecimento/Inteligência Artificial
2010	RAMALHO, Rogério, Ap. S.	Ciência da Informação	Organização e Representação do Conhecimento

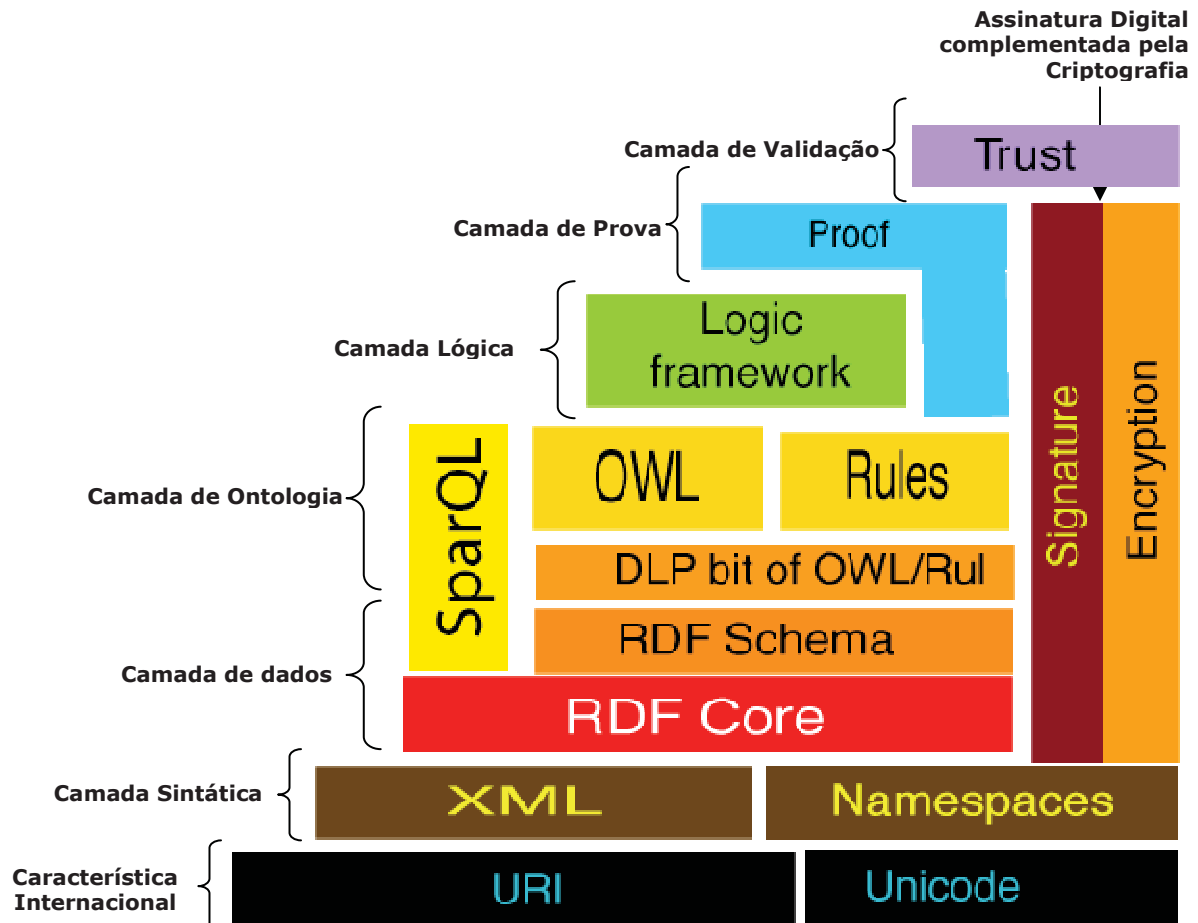
Fonte: Elaborado pelo autor.

No âmbito da *Web Semântica* exige-se que as informações tenham significado explicitados para que as máquinas possam automaticamente processar as informações encontradas na *Web*. Dessa forma, as ontologias atuam no fornecimento desse significado em camadas sobrepostas para a *Web Semântica*, para ser usado com *Extensible Markup Language* (XML), *XML Schema*, *RDF* e *RDF Schema*¹⁷, conforme pode ser observada na figura 3,

¹⁷ As tecnologias XML, *XML Schema*, *RDF* e *RDF Schema* são consideradas o núcleo para a efetivação da *Web Semântica* e para o estabelecimento da interoperabilidade entre ambientes informacionais digitais.

numa proposta de Berners-Lee (2005) para o delineamento de uma arquitetura da *Web Semântica*.

FIGURA 3: Arquitetura da *Web Semântica* comentada pelo autor.



Fonte: Adaptado de Berners-Lee (2005, p. 27).

Pode-se observar, de acordo com a figura 3, que a *Web Semântica* está dividida em camadas sobrepostas e para que ela funcione de forma efetiva, é necessário o trabalho conjunto e integrado das tecnologias que a constituem na sua totalidade para o seu potencial pleno. Nosso interesse de pesquisa está voltado principalmente para as camadas de representação e descrição dos dados. Dentre as camadas propostas por Berners-Lee (2005) destacam-se:

- **Característica Internacional:** composta por padrões de descrição e identificadores universais, tais como a URI (*Uniform Resource Identifier*) e UNICODE, que são responsáveis por uma identificação e localização mínima do recurso informacional na rede;

- **Camada Sintática:** responsável pelo estabelecimento da sintaxe de descrição dos dados, essa camada é composta pela linguagem XML, *namespaces*¹⁸ e também pelo XML *Schema*, que permite uma melhor estruturação dos dados e conseqüentemente uma melhor visualização do conteúdo dos recursos pelos agentes de *software*;
- **Camada de Dados:** essa camada trabalha com a representação, o processamento e a codificação dos metadados. É composta pela arquitetura de metadados e/ou modelo de dados RDF (RDF *Core*) – *Resource Description Framework* e o RDF *Schema*, que são ferramentas responsáveis por expressar significados e promover a interoperabilidade. É importante destacar que nesta camada e na camada anterior está implícito o uso de padrões de metadados para garantir a representação dos recursos informacionais;
- **Camada de Ontologia:** o uso de ontologias permite o estabelecimento da semântica, ou o significado dos dados descritos e representados pelos metadados. A definição dos conceitos nessa camada servirá para os agentes inteligentes “compreenderem” a semântica dos dados, e a linguagem recomendada pelo *World Wide Web Consortium* (W3C) é a OWL – *Ontology Web Language*. Vale destacar que a SparQL é uma linguagem computacional utilizada para realizar consultas sobre estruturas RDF, incluindo estruturas descritas por meio da linguagem OWL. Como não possui mecanismos de inferência, possibilita manipular apenas informações armazenadas em um meio físico qualquer, em outras palavras, define as relações entre os dados armazenados na camada de estrutura.
- **Camada Lógica:** é responsável por uma recuperação mais eficiente devido ao uso de agentes, regras (presentes também na camada de ontologia) e mecanismos de inferência, mas para isso contam como base as camadas responsáveis pela estruturação, representação e estabelecimento semântico dos dados, para posteriormente relacionar e processar as informações de forma lógica. Além disso, define regras lógicas que possam ser verificadas computacionalmente, permitindo a

¹⁸ *Namespace* pode ser compreendido como um conjunto de nomes identificados por uma referência URI ou um lugar onde reside um esquema (conjunto de nomes). Ele pode ser considerado com um único nome que identifica uma busca de onde o esquema originou-se. Por exemplo: “<http://purl.org/dc/elements/1.1/>” é um *namespace* para Dublin Core 1.1. (ZENG; QIN; BACA, 2008, p. 323, tradução nossa).

realização de inferências automáticas e a verificação do nível de coerência lógica dos recursos.

- **Camada de Prova:** responsável pelo intercâmbio entre agentes; é nessa camada onde serão processadas as definições lógicas pelos agentes para a construção da prova, que será criada para verificar se a dedução do agente estava correta;
- **Camada de Validação:** essa última camada é responsável pelo estabelecimento de verdades, autenticidade, confiabilidade e validade dos dados recuperados, garantindo aos agentes que raciocinam sobre os dados a veracidade da informação.

As ontologias são importantes se os “agentes” na *Web* são capazes de pesquisar e/ou mesclar informações de diversas comunidades. Isso ocorre porque o mesmo termo pode ser usado em diferentes contextos com diferentes significados, e o mesmo significado pode ser representado por termos diferentes em contextos distintos. A fim de padronizar os significados para a criação de ontologias que podem ser usadas na *Web*, a *Web Ontology Working Group*, do *World Wide Web Consortium* (W3C) desenvolveu a *Web Ontology Language* (OWL), ou seja, uma linguagem para a criação de ontologias interoperáveis que podem ser extensões de RDF.

Para Souza e Alvarenga (2004) as ontologias se apresentam como um modelo de relacionamento de entidades e suas interações, em algum domínio particular do conhecimento ou específico a alguma atividade. O objetivo de sua construção é a necessidade de um vocabulário compartilhado para se trocarem informações entre os membros de uma comunidade, sejam eles seres humanos ou agentes inteligentes.

No domínio bibliográfico, por exemplo, uma ontologia atuaria na definição dos conceitos dos atributos de um esquema de metadados, tal como as etiquetas do formato de intercâmbio MARC 21, onde o conteúdo, ou seja, o vocabulário poderia ser definido de acordo com as regras de descrição contempladas pelo código de catalogação vigente (AACR2).

Ontologias têm diferentes níveis de estrutura. Algumas ontologias são taxonomias isto é, classificações ou listas hierárquicas de terminologias de uma área particular. Algumas são

especificações de conjuntos de características conceituais. “Por exemplo, *metadada schemas* são ontologias que especificam elementos para ser usados, o que estes elementos significam, e quais tipos de atributos e valores aqueles elementos podem ter”. (TAYLOR, 2004, p. 284, tradução nossa). Outras ontologias aparecem para categorizar vocabulários controlados.

Uma ontologia tem por objetivo prover uma base semântica para esquemas de metadados, facilitando a comunicação entre sistemas e agentes. Assim, ontologias estabelecem fundamentos de significados conceituais sem os quais a *Web Semântica* não seria possível, devido à heterogeneidade dos conceitos representados. (JACOB, 2003, tradução nossa).

Para o campo da Ciência da Informação, as ontologias são uma área de pesquisa emergente, e seu interesse se fundamenta na potencialidade que elas têm em organizar e representar a informação, principalmente no que diz respeito ao melhoramento dos processos de recuperação da informação em múltiplos domínios e sua aplicação efetiva nos variados ambientes informacionais digitais.

Quando se exemplificam as ontologias como esquemas de metadados (TAYLOR, 2004), começa-se a perceber as sutis aproximações e suas correspondências, sobretudo, nas questões de representação de recursos¹⁹ e ainda os possíveis relacionamentos que só se efetivarão em um trabalho conjunto; pode-se dizer que há uma sinergia entre ambas, principalmente no delineamento e na modelagem de ambientes informacionais digitais e que poderão proporcionar a interoperabilidade semântica.

É sob essa orientação que se percebe a rica ligação entre ontologias e metadados, uma vez que para atender aos requisitos de interoperabilidade semântica entre padrões e/ou esquemas de metadados heterogêneos, é imprescindível uma base ontológica que suporte e que defina as variadas terminologias, como os campos, subcampos, indicadores e outros elementos de um registro bibliográfico, indispensáveis para uma representação e apresentação adequadas.

¹⁹ Nessa tese a expressão representação de recursos, refere-se à representação descritiva de recursos bibliográficos.

Nesse sentido, transportar semântica para domínios particulares é papel fundamental das ontologias, uma vez que são usadas amplamente com o intuito de representar o conhecimento de domínios específicos.

O uso comum de ontologias se dá na padronização e conceitualização dos dados através de uma linguagem formal de ontologia compreensível pela máquina/computador. Um esquema global em um sistema de integração de dados pode ser uma ontologia, que age como um mediador para unificar a heterogeneidade entre diferentes fontes de dados. (CRUZ; XIÃO, 2005).

Como um exemplo do uso de ontologias em rede para a integração de dados, seria vista na produção de um esquema fonte para uma ontologia local, que poderia ser acessível a outros pares, de modo a apoiar os mapeamentos semânticos entre diferentes ontologias locais. (CRUZ; XIÃO, 2005).

Ontologias têm sido amplamente utilizadas em sistemas de integração de dados, porque elas oferecem a explicitação de uma conceitualização de um domínio compreensível pela máquina/computador, podendo ser usada de três maneiras distintas, conforme destacam Cruz e Xião (2005):

- **Abordagem ontológica única:** todos os esquemas fontes estão diretamente relacionados com uma ontologia global compartilhada que fornece uma interface uniforme ao usuário. No entanto, essa abordagem requer que todas as fontes tenham claramente a mesma visão sobre um domínio, com o mesmo nível de granularidade.
- **Abordagens ontológicas múltiplas:** cada fonte de dado é descrita em seu próprio local separadamente. Em vez de usar uma ontologia comum, as ontologias locais são mapeadas para outras. Para essa proposta, um formalismo adicional na representação é necessário para definir os mapeamentos inter-ontologias.
- **Abordagem ontológica híbrida:** uma combinação de duas abordagens anteriores é usada. Primeiro, uma ontologia local é construída para cada esquema fonte, que, no entanto, não está mapeada para outras ontologias locais, mas para uma ontologia global compartilhada, Novas fontes podem ser facilmente adicionadas sem a

necessidade de modificar os mapeamentos existentes. A estrutura em camadas é um exemplo para esta abordagem.

Além disso, Cruz e Xião (2005) destacam a identificação de cinco utilizações das ontologias com o propósito de integração de dados:

- **Representação de metadados:** metadados (isto é, esquemas fonte) em cada fonte de dados podem ser explicitamente representados por uma ontologia local, utilizando uma linguagem única.
- **Conceitualização global:** a ontologia global fornece uma visão conceitual sobre esquemas fonte esquematicamente heterogêneos.
- **Suporte para consultas de alto nível:** dada uma visão de alto nível das fontes, tal como previsto por uma ontologia global, o usuário pode formular uma consulta sem conhecimento específico de fontes de dados diferentes. A consulta é então reescrita em consultas sobre as fontes, baseadas no mapeamento semântico entre ontologias globais e locais.
- **Mediação declarativa:** o processamento das consultas em sistemas híbridos *par para par* utiliza uma ontologia global com um mediador declarativo para reescrever consultas entre os pares.
- **Suporte ao mapeamento:** um tesouro, formalizado em termos de uma ontologia, pode ser usado para o processo de mapeamento a fim de facilitar a sua automação.

O mapeamento de esquemas de metadados às ontologias constitui-se em um processo difícil e complexo, uma vez que estas duas construções apresentam muitas diferenças em vários níveis. (SOWA, 2000; NISO, 2004).

Representar adequadamente um recurso informacional em quaisquer ambientes é uma condição para atender aos requisitos de qualidade e que refletirá nos processos de busca e de

recuperação pelo usuário final. E como dar condições aos agentes de *softwares* e/ou máquinas para encontrar e interpretar dados heterogêneos de sistemas distribuídos distintos?

Uma possível alternativa se baseia nos relacionamentos que são fundamentais para a semântica, no sentido de associar significado às palavras, aos itens e às entidades. Pode-se dizer que “os relacionamentos são considerados a chave para novas percepções. A descoberta de conhecimentos está sobre a descoberta de novos relacionamentos”. (SHET; ARPINAR; KHASYAP, 2002, p. 1, tradução nossa).

A *Web Semântica* visa associar anotações (isto é, metadados), com base primeiramente em conceitos (muitas vezes representando as entidades) de uma ou mais ontologias, com todos os recursos da *Web* acessíveis, tal que os programas possam associar o significado com os dados. (SHET; ARPINAR; KHASYAP (2002, p. 1, tradução nossa).

Contrapondo os autores Shet, Arpinar e Hhasyp (2002) ao afirmarem que os relacionamentos definem e possibilitam a compreensão por agentes computacionais interpretarem os dados de sistemas heterogêneos, se destaca primeiramente a construção padronizada de representações, com o uso de ontologias e de metadados, pautadas em normas, regras, esquemas e códigos internacionais, pois é através dessas e por meio de, se determinarão os relacionamentos e propiciarão a sua clareza, de uma forma mais efetiva, entre recursos e ambientes informacionais digitais heterogêneos e garantirão o uso e o (re)uso para os usuários do sistema.

É relevante destacar nesse momento as alterações entre metadados e ontologias, mas também apontar a necessidade de um olhar em um trabalho sinérgico entre tais instrumentos, no tocante às questões de interoperabilidade semântica entre ambientes informacionais digitais.

De acordo com as observações apresentadas sobre ontologias e o seu papel na interoperabilidade semântica, pode-se dizer que essas definem as entidades em um nível abstrato, com a intenção de conceitualizar um domínio de interesse particular. Elas não fornecem elementos específicos para a descrição de um recurso, mas uma visão geral de noções básicas de um campo de interesse e as relações entre essas noções.

Ontologias permitem uma melhor representação para os metadados de sistemas de integração: maior semântica, especialmente no que diz respeito às fontes e aos mapeamentos e, na integração de esquemas e/ou padrões de metadados. Enquanto os metadados são usados para descrever e representar recursos em termos de elementos, e facilitar a descoberta e o fácil acesso à informação.

2.2. Metadados: exigência para modelagem de ambientes informacionais digitais

Os bibliotecários produzem e padronizam metadados há séculos, desde as primeiras tentativas de organização da informação a partir da descrição de documentos. O que vem acontecendo ultimamente é que profissionais de diversas áreas estão buscando criar instrumentos de descrição da informação, mas seu desconhecimento dos métodos, processos e peculiaridades característicos da documentação, da Biblioteconomia, tem gerado uma variedade de padrões que muitas vezes não atende satisfatoriamente às exigências de uma lógica descritiva estabelecida e que dê conta da complexidade da caracterização desse material. (MILSTEAD; FELDMAN, 1999; ALVES, 2005; CASTRO; SANTOS, 2007; CASTRO, 2008).

Pode-se dizer que o objetivo e a função dos metadados estão fundamentados nos princípios da catalogação, ou seja, garantir a padronização dos recursos informacionais (forma e conteúdo), pautados em normas e regras internacionais na tentativa de facilitar e potencializar a identificação, a busca, a localização, a recuperação, a preservação, o uso e o (re) uso dos recursos informacionais. A diferença dessa forma de representação está na nova abordagem dada pelo ambiente tecnológico em que ela se insere. (ALVES, 2005; CASTRO, 2008).

Metadados não podem ser descrições desestruturadas de recursos. Eles devem ser padronizados e controlados. Sem regras formais, normas e padrões, a descrição de metadados não é melhor do que o acesso aos recursos informacionais por palavras-chave.

Não é objetivo nessa tese tratar as conceituações, as características, os tipos, as classificações de metadados, esquemas e padrões de metadados, uma vez que é possível

encontrar investigações dessa natureza nos trabalhos de Castro e Santos (2007), Castro (2008) Zeng e Qin (2008), Baca (2008), Castro e Santos (2009), Santos e Alves (2009), Castro e Santos (2010) e Alves (2010), dentre outros.

Pontua-se que essa investigação estará embasada na definição de metadados e padrões de metadados de Alves (2010, p. 47-48, grifo do autor), pois esta se aplica aos propósitos dessa tese e permeia a construção de ambientes informacionais digitais melhor estruturados.

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

Os **padrões de metadados** são estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados. O objetivo do padrão de metadados é descrever uma entidade gerando uma representação unívoca e padronizada que possa ser utilizada para recuperação da mesma. (ALVES, 2010, p. 47-8).

As unidades básicas de metadados são os esquemas e os elementos. *Elementos de metadados* são as categorias individuais ou classes que mantêm ou sustentam as partes individuais da descrição de um conjunto/pacote de informação. Elementos de metadados típicos incluem, por exemplo, título, criador, data de criação, identificação de assunto e a preferência. Na visão de Zeng e Qin (2008, p. 321, tradução nossa) o elemento de metadado (*Metadata element*) “é um termo formalmente definido para descrever uma das propriedades de um recurso de um tipo particular ou para uma proposta particular. Por exemplo, o ‘publicador’ de um livro, ‘o formato’ de um arquivo eletrônico, ou uma ‘restauração de dados’ de uma construção”.

Enquanto isso, os *Esquemas de metadados* são conjuntos de elementos desenhados para encontrar as necessidades de comunidades particulares. Enquanto alguns esquemas são gerais em sua natureza, têm sido criados mais esquemas para tipos de informação específica. “Esquemas têm sido desenhados para lidar com informação governamental, geoespacial, busca visual e muitos outros tipos de pacotes/conjuntos de informações”. (TAYLOR, 2004, p. 147, tradução nossa). Esquemas podem variar grandemente, por exemplo, no número de elementos de dados, no uso de elementos obrigatórios e repetíveis, em codificação e no uso de

vocabulários controlados, entre outras coisas. Enquanto mais esquemas focam em elementos descritivos para suportar/sustentar a descoberta do recurso, alguns contêm elementos para suportar propostas administrativas e estruturais. “Com as várias necessidades de diferentes comunidades, não é possível criar um esquema perfeito, que contempla a totalidade do metadado”. (TAYLOR, 2004, p. 145, tradução nossa).

De acordo com Zeng e Qin (2008, p. 323, tradução nossa) o esquema de metadados (*Metadata schema*) pode ser considerado como:

Uma especificação processavelmente por máquinas que define a estrutura, a codificação de sintaxe, regras, e formatos para um conjunto de elementos de metadados em uma linguagem formal num esquema. Na literatura o termo “*metadata schema*” usualmente refere-se ao conjunto de elementos na sua totalidade, assim como a codificação dos elementos e a estrutura com uma linguagem de marcação.

De acordo com Vellucci (2000, tradução nossa) há três características encontradas em todos os esquemas de metadados: (1) estrutura (2) sintaxe (3) semântica.

- (1) Estrutura: refere-se ao modelo de dados ou arquitetura usada para manter/comportar o metadado e a forma como as declarações (*statements*) do metadado são expressas. Estrutura aqui é entendida ou está sendo referida à estrutura do metadado. Não deve ser confundida com “metadado estrutural”, que se refere à estrutura descrita inicialmente do recurso. Dois exemplos de tais modelos são RDF e METS (*Metadata Encoding and Transmission Standard*).
- (2) Sintaxe: refere-se à codificação do metadado. Este pode ser o formato MARC (*Machine Readable Cataloging*) para registros bibliográficos ou um XML ou SGML DTD²⁰ (*Standard Generalized Markup Language Document Type Definition*) para outros tipos de metadados.
- (3) Semântica: refere-se ao significado, especialmente o significado de vários elementos de dados. A semântica ajuda aos criadores a entender, por exemplo, o que “cobertura”

²⁰ DTD (*Document Type Definition*) – estabelece um conjunto de regras (sintaxe) para a validação da estrutura de um documento XML. Uma das desvantagens da DTD é que ela permite apenas a verificação sintática de um documento sem nenhum tipo de controle semântico (conteúdo). “[...] pode-se garantir que o título foi digitado, mas não se realmente era o título, ou ainda, pode-se verificar se o dado informado no campo ‘data de publicação’ é realmente uma data, mas não se garante que seja a data correta”. (SIQUEIRA, 2003, p. 75).

ou “modificação de data” significa em um dado esquema. A semântica de um esquema de metadados não dita o espaço do conteúdo junto aos elementos. Esta é a “província” de padrões de conteúdo (ou regras de conteúdo) e vocabulários controlados. Os padrões de conteúdo determinam tais coisas, assim como a data será formatada junto aos elementos do metadado. Por exemplo, ele deve especificar que todas as datas serão entradas usando o formato YYYY-MM-DD, sendo YYYY o ano; MM o mês; e DD o dia. Vocabulários controlados se referem a listas de palavras os quais certos termos são escolhidos e seus sinônimos, limitando a extensão de valores que podem ser inseridos dentro de uma classe. Vocabulários controlados são frequentemente usados em tipos de objetos e assuntos relacionais de elementos de dados. Se tais regras e sistemas não existissem, a recuperação da informação de uma forma efetiva poderia ser comprometida.

Para exemplificar as características encontradas num esquema de metadados, tem-se a figura 4, com o atributo (245) do formato de intercâmbio bibliográfico MARC 21. O campo, etiqueta, metadado ou atributo²¹ (245) equivale ao título do recurso bibliográfico e possui os subcampos (a) informação sobre o título do recurso; (b) subtítulo do recurso; e (c) indicação de responsabilidade sobre o recurso bibliográfico. Os campos e subcampos no esquema MARC 21 correspondem à sintaxe do registro bibliográfico, e precisam ser preenchidos de acordo com as regras de conteúdo (AACR2) que definem e padronizam a descrição e a representação bibliográfica (semântica).

FIGURA 4: Etiqueta título do formato MARC 21.

245 00 | a Introduction to metadata : | b pathways to digital information / | c edited by Murtha Baca.

Fonte: elaborado pelo autor.

De acordo com as características dos esquemas de metadados, Zeng e Qin (2008, p. 93, tradução nossa) apontam que dentre elas, duas não podem ser esquecidas, consideradas componentes básicos no conjunto de elementos do esquema, as quais se destacam: “[...] Semântica (definições dos significados dos elementos e seus relacionamentos); e Conteúdo (declarações ou instruções do que e como os valores devem ser atribuídos aos elementos)”.

²¹ Principais terminologias encontradas na literatura de Catalogação e Metadados para caracterizar a constituição de um padrão e/ou esquema de metadados.

No que concerne à semântica, Zeng e Qin (2008) apontam que os esquemas de metadados são criados para a identificação e a descrição de recursos e não expressam semântica rica²². Mesmo que o significado das informações dos metadados e seus relacionamentos para descrever recursos podem ser entendidos e processados por humanos, o processamento por máquinas e seus relacionamentos não é óbvio, ao menos que declarado. Ao contrário, as classes de ontologias são interligadas por propriedades específicas que declaram explicitamente os relacionamentos semânticos entre essas entidades. Mapeando um esquema de metadado real para uma ontologia gera muitas dificuldades devido ao grande número de expressões conceituais que devem ser alinhados.

Quando um conjunto de elementos de metadados é codificado, os vários resultados são chamados de “esquemas de especificações processáveis por máquinas” que definem a estrutura e a sintaxe de metadados em uma linguagem de esquema formal. DCMI tem utilizado tanto XML e RDFs para apoiar o Dublin Core simples (sem refinamentos) e Dublin Core qualificado (com elementos de refinamento e esquemas de codificação). (ZENG; QIN, 2008).

Pode-se dizer que bons modelos de estruturas é resultado do uso correto de princípios, os quais podem ser encontrados e há muito tempo consolidados nas metodologias cristalizadas e sedimentadas das áreas de Biblioteconomia e Ciência da Informação, por meio do uso dos códigos de catalogação, por exemplo, os quais padronizam a elaboração e a construção dos metadados e dos esquemas de metadados nos variados ambientes informacionais (digitais ou não).

Independentemente da arquitetura dos sistemas de integração de dados, os metadados exercem um papel fundamental nesses sistemas, por permitirem descrever não só os dados armazenados pelas fontes, mas também diversos outros aspectos de sistemas de integração de dados, tais como mapeamentos, esquemas, entre outros. (TAYLOR, 2004; ZENG; QIN, 2008).

²² A expressão rica (o) refere-se às estruturas com maior número de elementos de descrição e seus dados são codificados por algum esquema de descrição bibliográfica. Exemplo: Estrutura MARC 21 e esquema AACR2.

Em relação ao padrão de metadados, pode-se destacar o formato de intercâmbio MARC 21, do domínio bibliográfico, apontado por apresentar e construir representações completas e complexas do recurso bibliográfico.

2.2.1 Padrão de metadados MARC 21: complexidade na representação e descrição bibliográfica

Os processos de catalogação foram evoluindo, ao longo do tempo, e nota-se em seu caminho a utilização das tecnologias disponíveis no intuito de proporcionar uma padronização na representação e ao mesmo tempo reduzir custos, tempo e principalmente possibilitar o compartilhamento de informações. Um exemplo da evolução da catalogação foi a criação na década de 1960 do formato MARC – *Machine Readable Cataloging* – que possibilitou não só a automação das bibliotecas, como também possibilitou o intercâmbio de registros bibliográficos legíveis por máquina. Hoje o formato MARC se constitui como um formato ou padrão de representação informacional já consolidado e muito utilizado na área de Biblioteconomia.

O MARC pode ser definido como um “[...] formato padronizado para armazenamento e intercâmbio de registros bibliográficos e informações relacionadas em formato legível por máquina” (BRITISH LIBRARY, 2003). De acordo com Alves (2005, p. 141),

Esse formato é destinado à criação, armazenamento, gerenciamento e intercâmbio de registros catalográficos e bibliográficos. O objetivo geral do MARC 21 é fornecer aos usuários a localização, o acesso e a recuperação dos recursos informacionais por meio de uma representação padronizada, e possibilitar a comunidade que adota esse formato intercambiar registros bibliográficos, ou seja, intercambiar os registros com a representação dos recursos informacionais.

Após a sua criação, o formato MARC passou a ser utilizado em diversos países e com o tempo sofreu adaptações, dando origem a estruturas de representação diferenciadas, de acordo com a realidade de cada localidade e ao mesmo tempo uma incompatibilidade de formatos. Assim, na tentativa de minimizar esforços na catalogação e intercâmbio de dados buscou-se uma unificação dos diversos formatos MARC gerados no intuito de criar uma versão única e definitiva. Essa versão passou a ser denominada MARC 21 (formato MARC

para o século 21) e com ele foi possível evitar o trabalho duplicado das instituições e conseqüentemente estabelecer um maior intercâmbio dos registros bibliográficos (SIQUEIRA, 2003). Não serão tratadas aqui as alterações pela qual passaram as versões do MARC, o intuito é apenas proporcionar um breve panorama de sua criação e aprofundar nas questões de sua estrutura descritiva.

O formato MARC 21 é um padrão composto por um conjunto de cinco formatos coordenados que se apresenta subdividido da seguinte forma:

- Formato MARC 21 para Dados Bibliográficos: contém especificações para descrição de informações bibliográficas sobre diversos tipos de materiais ou recursos informacionais (materiais textuais impressos e manuscritos, arquivos de computador, mapas, músicas, recursos contínuos, materiais visuais entre outros);
- Formato MARC 21 para dados de Autoridade: contém especificações para descrição de dados de autoridade (nome do responsável pela obra) e descrição de assunto (vocabulários controlados e listas de cabeçalho de assunto) a serem utilizados como pontos de acesso no registro bibliográfico;
- Formato MARC 21 para controle de dados: contém especificações para a codificação de elementos que são identificados por um código no Líder (posição 06), informação utilizada para o gerenciamento de recursos.
- Formato MARC 21 para dados de classificação: contém especificações para a codificação de elementos relacionados a esquemas de classificação.
- Formato MARC 21 para comunidade informacional: contém especificação para a codificação de informações sobre recursos não bibliográficos destinados a uma comunidade específica. (LIBRARY OF CONGRESS, 2006).

Embora as características do formato MARC 21 sejam comuns a todos os formatos pertencentes à família MARC será abordado aqui mais detalhadamente o formato MARC 21 para Dados Bibliográficos.

A estrutura do formato MARC 21 é composta por três elementos: estrutura do registro, indicação de conteúdo e conteúdo dos elementos que compõe o registro. A seguir apresenta-se cada um deles, com suas respectivas especificações.

a) Estrutura do registro: baseada em norma internacional da *American National Standard for Bibliographic Information Interchange* - ANSI / NISO Z39.2 - para intercâmbio de informação bibliográfica e sua norma ISO equivalente – *Format for Bibliographic Information Interchange on Magnetic Tape* - ISO 2709 (FERREIRA, 2002; FURRIE, 2000).

b) Indicação ou designação de conteúdo: são códigos e convenções definidos no próprio MARC 21 para identificar e caracterizar os dados dentro do registro e ainda permitir sua manipulação, esses códigos são as etiquetas, indicadores, subcampos etc. (FERREIRA, 2002; FURRIE, 2000). Deste modo, pode-se dizer que a indicação do conteúdo está estruturada da seguinte forma, conforme aponta Furrie (2000):

- Campo: são unidades lógicas que contêm informações que representam o recurso informacional. Cada campo representa um tipo de informação bibliográfica do recurso, tal como o campo que representa o título e indicação de responsabilidade;
- Etiqueta: associada ao campo a etiqueta é composta por um número de três dígitos que identifica o campo e o dado (representação) que se segue. Algumas podem ser repetidas (duplicadas) outras não. Um exemplo que poder ser dado é a etiqueta 245 que representa e define o campo de título e indicação de responsabilidade.
- Indicadores: são as duas primeiras posições de caracteres que se seguem depois da etiqueta, a partir do campo 010. Sua função é definir com maior detalhe os dados descritos nos campos para serem interpretados pelas máquinas, porém, não são utilizados em todos os campos. Os indicadores podem conter valores de 0 a 9 e apesar de aparecerem juntos são números individuais que representam separadamente uma característica dentro do registro bibliográfico. Um exemplo de seu uso pode ser dado sobre o campo 245 com indicador 1 igual a 1, que significa entrada secundária de título e indicador 2 igual a 0 que significa o número de caracteres a desprezar na busca/alfabetação (caso o título do recurso não comece por nenhum artigo ou palavra que deva ser descartada na alfabetação).
- Subcampo: alguns campos são subdivididos em subcampos que são elementos de informação que representam certos tipos de dados dentro de um campo e alguns deles

podem ser repetidos (duplicados). Os subcampos são compostos por códigos de subcampos, conforme será visto a seguir.

- Código de subcampo: os subcampos são representados por códigos de subcampos, que se assemelham a subetiquetas, e são formados por uma letra minúscula antecedida de um delimitador (símbolo). Cada letra utilizada para a indicação do subcampo representa uma informação diferente a ser representada.
- Designador de conteúdo: essa denominação é utilizada para referir-se ao conjunto de etiquetas, indicadores e códigos de subcampo que, segundo Furrie (2000) se constituem como chave do sistema de notação MARC 21. Para este autor, “Os três tipos de designadores de conteúdo são os símbolos taquigráficos que marcam e explicam o conteúdo de um registro bibliográfico” (FURRIE, 2000). O designador de conteúdo pode ser considerado como um elemento complexo de representação na estrutura do formato MARC 21.

c) Dados contidos no registro: o conteúdo dos elementos do registro MARC é composto pela representação de informações, em outras palavras, pelos metadados que descrevem o recurso informacional. Essas informações geralmente são definidas por padrões externos, tais como ISBD (*International Standard Bibliographic Description*), AACR2 (*Anglo American Cataloging Rules – segunda edição*), LCSH (*Library of Congress Subject Headings*) etc., com exceção do líder e campos 007 e 008 que são definidos no próprio MARC (FERREIRA, 2002; FURRIE, 2000).

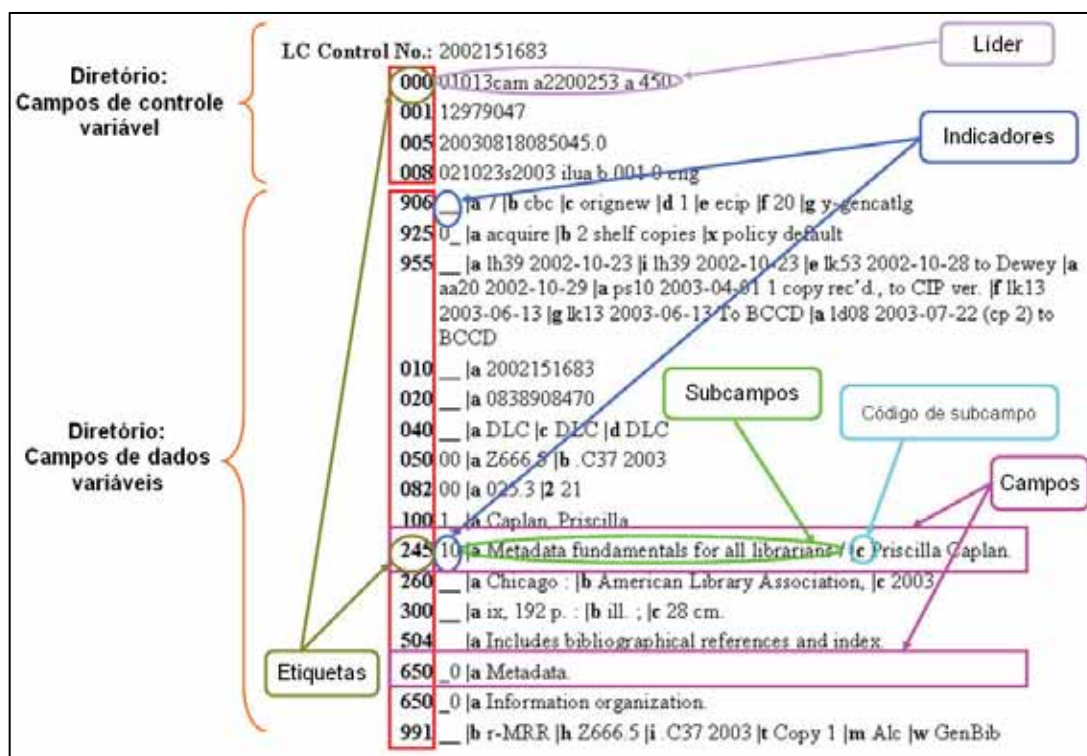
Dois desses três elementos, componentes do registro bibliográfico merecem maior destaque, pois atuam diretamente na construção de representações por metadados; são eles: os dados contidos no registro, definidos por regras externas e a indicação ou designação de conteúdo, que ainda comporta uma outra especificação de estrutura, conforme aponta Ferreira (2002).

Em relação à composição do registro MARC 21, vale dizer que é formado por três elementos principais:

- Líder: primeiro campo do registro bibliográfico que fornece informações para o processamento do próprio registro. É constituído por 24 caracteres que são gerados automaticamente, representando um resumo do registro bibliográfico;
- Diretório: uma série de entradas do registro bibliográfico, que contém a posição inicial e o tamanho de cada etiqueta dentro do registro de identificação. É composto primeiramente pelos campos de controle variável, seguidos dos campos de dados variáveis (arranjados em ordem crescente de acordo com o primeiro caractere da etiqueta);
- Campos Variáveis: os dados em um registro bibliográfico MARC 21, estão organizados em campos variáveis, sendo cada campo identificado por uma etiqueta registrada na entrada do diretório. Os campos variáveis estão subdivididos em dois tipos:
 - Campos de controle variável: são os campos 00X, identificados por etiquetas que não contém posições nem para indicadores nem para códigos de subcampo. Os campos de controle variável podem conter um único dado ou uma série de dados de tamanho fixo, identificados pela posição relativa do caractere.
 - Campos de dados variáveis: os outros campos restantes são do tipo variáveis, também são representados por etiquetas, mas se diferem dos anteriores por conter duas posições para indicadores, localizadas no começo de cada campo, e dois caracteres para código de subcampo, precedendo cada dado dentro do campo. Os campos de dados variáveis são agrupados em blocos, de acordo com o primeiro caractere da etiqueta, o qual, com algumas exceções, identificam a função do dado dentro do registro. O tipo de informação no campo é identificado pelo restante da etiqueta.

Os elementos definidos anteriormente podem ser visualizados na figura 5.

FIGURA 5: Estrutura do formato MARC 21.



Fonte: Library of Congress (2011).

A estrutura do formato MARC 21 possibilita a descrição completa do recurso informacional de todos os tipos, incluindo recursos disponíveis em meio eletrônico e sua posterior interoperabilidade entre sistemas. Por ser baseado em regras de catalogação, requer o uso de esquemas externos de representação, tal como o AACR2, LCSH, entre outros, caracterizando-se por possuir uma estrutura de representação lógica e coerente que estabelece uma relação semântica entre os elementos descritos.

De acordo com Siqueira (2003) a maior crítica sobre o MARC 21 recai sobre a sua complexidade, que é justificada pelo fato da estrutura do padrão apresentar um alto grau de detalhamento, pois foi construído para representar qualquer tipo de recurso informacional do modo mais detalhado e completo possível. Outra questão apontada está relacionada ao uso de esquemas, tal como o uso “obrigatório” do AACR2 ou das ISBDs. Sobre este aspecto Siqueira (2003, p. 66) afirma que,

Lembramos que se tratam de dois instrumentos distintos, o AACR2 é um código de catalogação para descrição de dados bibliográficos e catalográficos e o MARC 21 é um formato de intercâmbio de dados bibliográficos e catalográficos, entretanto, visto que a principal preocupação da catalogação como forma de representação é a padronização e a qualidade

dos registros, pode-se afirmar que a utilização conjunta de ambos é justificável.

O formato MARC 21 se encontra na versão em XML – *Extensible Markup Language* – mantendo todas as características do formato MARC 21 original, mas com as vantagens da XML, uma estrutura mais flexível, hierárquica e organizada.

A linguagem XML permitirá uma maior interoperabilidade dos dados e permitirá um maior gerenciamento nas formas de representação disponíveis. Assim, o MARC XML agrega não só as vantagens de uma representação padronizada e completa dos recursos informacionais, como também a flexibilidade da linguagem XML garantindo uma maior interoperabilidade entre sistemas e ambientes informacionais que trabalham com metadados diferentes (SIQUEIRA, 2003).

De acordo com Miller (1996) cada padrão de metadados fornecerá uma representação diferente e quanto mais completa e detalhada for essa representação melhor será a recuperação dos recursos informacionais.

É sobre esse ponto de vista que defende-se o formato MARC 21 como o padrão de metadados mais adequado para a representação de informações em ambientes informacionais digitais. Desse modo, o MARC apresenta as seguintes vantagens: proporciona uma representação completa e detalhada de qualquer tipo de recurso informacional, incluindo recursos em meio digital; possui regras externas (esquemas) que norteiam a construção das representações em sua estrutura de descrição, garantindo maior clareza, padronização e confiabilidade da representação e possibilita o intercâmbio de registros bibliográficos entre instituições e a interoperabilidade entre sistemas.

Independente da versão utilizada, MARC 21 ou MARC XML, o que norteia a sua utilização é o princípio de criação do padrão. Segundo as definições abordadas sobre o MARC pode-se dizer que o princípio de criação do formato está pautado na necessidade de criar uma estrutura padronizada de descrição para a criação, o armazenamento, o gerenciamento, a localização e o intercâmbio de representações bibliográficas de um recurso informacional.

Com a heterogeneidade dos recursos informacionais, armazenados e estruturados nos mais diversos esquemas, a literatura científica aponta algumas técnicas de mapeamento de elementos de metadados, entre padrões de metadados distintos. Nessa tese, nos deteremos especialmente ao *crosswalk*, pois é esta uma ferramenta tecnológica que contribui para os aspectos de interoperabilidade em ambientes informacionais digitais.

2.3 *Crosswalk*: mapeamento entre esquemas de metadados

A *National Information Standards Organization* (NISO, 2004) define um *crosswalk* como um “mapeamento de elementos, semântica e sintaxe de um esquema de metadados para outros”.

Crosswalks é um método muito utilizado para permitir a interoperabilidade entre esquemas de metadados. O método começa com esquemas de metadados independentes; tentativas são feitas para mapear ou criar *crosswalks* entre termos de metadados equivalentes ou comparáveis, isto é, elementos e refinamentos. (Algumas vezes outros termos são usados para se referir ao elemento, por exemplo, campo, nível e etiqueta). O mecanismo usado em *crosswalk* é normalmente um gráfico ou tabela que representa o mapeamento semântico dos elementos de dados em outro padrão de dados (fonte/origem) para aqueles em outro padrão (alvo) baseado na similaridade da função ou significado dos elementos. (BACA, 2000-2008; ZENG; QIN, 2008).

O *Crosswalks* permite aos sistemas efetivamente converter dados de um padrão de metadados para outro. Permite também coleções heterogêneas ser pesquisadas simultaneamente com uma única consulta, como se contido em um único banco de dados (interoperabilidade semântica).

Nos últimos anos, grandes esforços no mapeamento de metadados têm produzido um número considerável de *crosswalks*. Quase todos os esquemas criaram *crosswalks* para esquemas populares, tais como, Dublin Core e MARC. Especificações de metadados podem também incluir *crosswalks* para uma versão prévia, bem como para outros esquemas. (BACA, 2000-2008; ZENG; QIN, 2008).

O método predominante usado em *crosswalking* é o mapeamento direto ou o estabelecimento da equivalência de elementos entre esquemas diferentes. O mapeamento de metadados se refere à identificação formal de equivalência ou quase equivalência de elementos de metadados ou grupos de elementos de metadados de diferentes esquemas de metadados, realizados para facilitar a interoperabilidade semântica. (BACA, 2008).

Muitas propriedades de metadados precisam ser levadas em consideração para o mapeamento. De acordo com o documento da NISO *Issues in crosswalking content metadata standards* (St. Pierre and LaPlant, 1998), as propriedades comuns podem incluir:

- Uma definição semântica de cada elemento de metadado;
- Se um elemento de metadado é obrigatório, opcional ou obrigatório baseado em certas condições;
- Se um elemento de metadado pode ocorrer várias vezes no mesmo registro;
- As restrições impostas à organização dos elementos de metadados em relação ao outro, por exemplo, os relacionamentos hierárquicos pai-filho;
- As restrições impostas sobre o valor de um elemento (ex., texto livre, intervalo numérico, data, ou vocabulário controlado) e;
- Suporte opcional para elementos de metadados definidos localmente. (ZENG; QIN, 2008, p. 120, tradução nossa).

Baca (2008), Zeng e Qin (2008) elucidam que duas abordagens têm sido utilizadas na prática do *crosswalking*:

A abordagem *absolute crosswalking* requer um mapeamento exato entre os elementos envolvidos (por exemplo, *vra.title* → *dc.title*) de um esquema de origem (exemplo, VRA Core²³) e um esquema de destino (exemplo, DC). Quando não existe equivalência exata, não haverá *crosswalking* (exemplo, *vra.technique* → [espaços vazios]). *Absolute crosswalking* assegura a equivalência (ou estreitamente equivalentes encontrados) dos elementos, mas não funciona bem para a conversão de dados.

²³ VRA Core é um padrão de metadados para recursos visuais que foi desenvolvido pela Visual Resources Association's Data Standards Committee. É constituído por um conjunto de elementos de metadados (tais como título, local, data etc.) juntamente com um plano inicial de como os elementos podem ser estruturados hierarquicamente. O conjunto de elementos oferece uma organização categórica para a descrição de obras de cultura visual, bem como as imagens que os documentam. (VRA, 2010, tradução nossa).

O problema nessa abordagem é que os valores dos dados em espaços não autorizados/qualificados são deixados de fora, especialmente quando um esquema de origem tem uma estrutura mais rica do que o esquema de destino. Um exemplo dessa situação pode ser vista na transformação do formato MARC 21, um padrão de metadados rico, para o padrão de metadados Dublin Core (simples). Nesse caso específico ocorre à perda de dados no processo, numa leitura e/ou compreensão e interpretação dos esquemas pela máquina, o que numa representação futura, na camada de apresentação das informações para os usuários gerará a inconsistência de dados, e ainda poderá comprometer a recuperação de informações significativas e relevantes.

Para solucionar esse problema, Baca (2008) apresenta uma resolução alternativa, o *relative crosswalking*, usado para mapear todos os elementos em um esquema de origem para pelos menos um elemento de um esquema de destino, independentemente se os dois elementos são equivalentes semanticamente (ex.: *vra.technique* → *dc.format*). A abordagem *relative crosswalking* parece um mapeamento que funciona melhor de um esquema complexo para um simples (ex.: de MARC para Dublin Core), mas não vice-versa.

Um dos problemas de *crosswalking* são os diferentes níveis de equivalência: um-para-um, um-para-muitos, muitos para um, e um para nenhum/nada. (BACA, 2008). Estas situações ocorrem em muitos cruzamentos de metadados, especialmente mapeamentos entre uma estrutura horizontal (ex.: Dublin Core) e uma estrutura hierárquica (ex.: MODS²⁴); e entre duas estruturas hierárquicas que são significativamente diferentes. O nível de detalhes pode estender somente de elementos simples para elementos mais qualificados/refinados ou subelementos.

No entanto, geralmente apenas os nomes dos elementos e suas definições são levados em consideração em um *crosswalk*. Isso significa que quando se mapeiam elementos individuais, muitas vezes não há equivalência/correspondência exata. Ao mesmo tempo, muitos elementos encontram-se à sobreposição no significado e no escopo/alcance. Por estes motivos, a conversão de dados baseada em *crosswalks* gera problemas de qualidade. (BACA, 2000-2008; ZENG; QIN, 2008).

²⁴ *Metadata Object Description Schema* (MODS) é um esquema para um conjunto de elementos bibliográficos que podem ser utilizados para diversas propostas, e particularmente para aplicações em bibliotecas. O padrão é mantido pela *Network Development e MARC Standards Office da Library of Congress* (LC). (MODS, 2010).

Um esquema de intercâmbio (novo ou já existente) para o canal *crosswalking* entre múltiplos esquemas, tornou-se uma solução bem aceita. Neste modelo, um dos esquemas é usado como mecanismo de troca entre vários esquemas. Em vez do mapeamento entre cada par ou grupo, cada um dos esquemas individuais de metadados é mapeado apenas para o esquema de intercâmbio. (BACA, 2008).

A figura 6 apresenta um exemplo do mapeamento de metadados pelo processo de *crosswalk* entre os padrões Dublin Core e MARC 21.

FIGURA 6: Exemplo de mapeamento entre os elementos de DC e MARC 21.

Elemento	Dublin Core	MARC 21
Autor	Creator	700 1#\$a (Adicionar entrada nome pessoal com \$e=autor) 720 \$a (Adicionar entrada- nome não controlado, com \$e=autor)
Título	Title	245 00\$a (Título principal)
Data	Date Create	260 ##\$c (Data de publicação, distribuição etc.).

Fonte: Elaborado pelo autor.

Percebe-se pela figura 6, a correspondência entre os elementos de padrões de metadados heterogêneos, por exemplo, *Title* (DC) e o campo 245 do formato MARC 21 (Título principal), ambos os elementos têm como semântica a informação do título de um recurso informacional ou de um registro bibliográfico.

Crosswalks têm um impacto direto sobre a qualidade da conversão de dados: qualquer correspondência ou falha nos pares (esquemas) pode causar conversão incorreta ou perda de milhares de valores, em outras palavras, gerar a inconsistência dos dados tanto na camada de descrição e representação, quanto na camada de apresentação dos dados num sistema de busca e recuperação de informações.

Além da técnica de *crosswalk*, fato importante a ser considerado nos metadados e esquemas de metadados, são as arquiteturas de metadados, ou seja, uma estrutura que permita comportar os metadados, de forma que tanto a sintaxe, quanto à semântica possam ser compreendidas por uma máquina, por exemplo, e ainda propiciar a interoperabilidade semântica entre ambientes e sistemas de informação heterogêneos.

2.4 Arquitetura de metadados: requisitos para a interoperabilidade

Num cenário marcado pela explosão informacional torna-se imprescindível que os ambientes informacionais atuais e altamente distribuídos, como a *Web*, compartilhem e colaborem com informações organizadas e estruturadas. Tal realidade tem desafiado profissionais de diversas áreas do conhecimento, sobretudo, os da área da Ciência da Informação a buscar soluções para o tratamento informacional de tais ambientes, a fim de que possa ocorrer a comunicação eficiente entre estes.

Dentre as iniciativas para promover que os ambientes informacionais possam comunicar-se entre si, um dos requisitos está na adoção de arquiteturas de metadados, ou seja, uma estrutura que permita integrar e alocar diferentes tipos de padrões metadados na *Web*, ou seja, garantir a interoperabilidade.

As arquiteturas de metadados têm como propósito a representação, a codificação e o transporte de padrões e/ou esquemas de metadados heterogêneos com estruturas flexíveis, com o objetivo de promover a interoperabilidade nos níveis estrutural, sintático e semântico, fazendo com que os recursos informacionais nos ambientes digitais possam estar cada vez mais e melhores representados e descritos.

Com a heterogeneidade de dados e metadados, disponibilizados no ambiente *Web*, faz-se necessário o desenvolvimento de um esquema ou estrutura que comporte esta diversidade de metadados, de modo que permita sistemas interoperáveis na rede. Esse esquema ou estrutura é denominado arquitetura de metadados, cujos objetivos são representar e proporcionar uma base de codificação de metadados por meio de estruturas flexíveis.

Destaca-se que essa tese tem como ponto focal a arquitetura de metadados RDF *Resource Description Framework*, que será tratada no capítulo 4, pois ela é considerada o pilar da descrição e a representação bibliográfica semântica, no mapeamento de padrões e/ou esquemas de metadados distintos e tem sido apontada atualmente como a solução que garantirá a interoperabilidade nos ambientes informacionais digitais, principalmente no domínio bibliográfico.

2.5 Integração estratégica entre ontologias e metadados

As ontologias e os metadados encontram-se como bases estruturantes das descrições bibliográficas atuais e podem ser consideradas como elementos básicos e norteadores para o estabelecimento da interoperabilidade semântica entre os ambientes e sistemas informacionais heterogêneos.

Nesse sentido, pode-se corroborar com a assertiva de Jacob (2003) ao se referir que todos os esquemas de metadados estão atrelados à ontologia, onde procuram especificar um conjunto de elementos (características físicas e/ou conceituais) de um recurso, julgadas relevantes para uma determinada comunidade ou área de conhecimento particular.

Reforçando a correlação entre ontologias e metadados, Greenberg, Sutton e Campbell (2003) apontam ainda que as ontologias são compreendidas como sistemas de metadados (ou vocabulários de metadados, constituindo na visão dos autores “a artéria central de metadados para a *Web Semântica*”.

Faz-se necessário uma rearticulação dos conceitos e bases epistemológicas que abarcam as ontologias e os metadados, na tentativa de inseri-los em domínios específicos de aplicação, como a Ciência da Informação, ou sua utilização para fins mais abrangentes, potencializando suas funcionalidades e permeadas pela dinâmica da evolução das novas tecnologias de informação e comunicação (TIC).

Retoma-se a idéia de que a heterogeneidade de recursos disponibilizados no ambiente *Web* atualmente requer o uso estratégico e efetivo do instrumental que pode contribuir para o

estabelecimento e para a modelagem dos catálogos de forma mais contundente, tais como as ontologias e os metadados, proporcionando aos sistemas de informação a idéia elementar de melhores formas de encontrar, identificar, selecionar e obter recursos informacionais e/ou registros bibliográficos no atendimento às necessidades dos usuários.

As ontologias e os metadados utilizados de forma estratégica e sinérgica poderão garantir a integridade das informações, fornecendo elementos de descrição mais ricos semanticamente e proporcionando ainda os relacionamentos entre um conjunto de esquemas, oriundos de ambientes informacionais heterogêneos, garantindo à interoperabilidade semântica e multidimensionando as formas de acesso aos recursos, para posterior utilização e reutilização de informações.

Dessa forma, é possível observar contemporaneamente o trabalho conjunto almejando a interoperabilidade semântica, no investimento nas camadas estruturantes e determinantes dos ambientes informacionais digitais, dadas pela representação e descrição, principalmente com a utilização de esquemas de metadados e ontologias que potencializam os processos de busca e de recuperação da informação de forma mais efetiva. Assim, o próximo capítulo apresenta a iniciativa *MarcOnt*, uma ontologia para a descrição bibliográfica, pensada no domínio bibliográfico aspirando ser um modelo referencial para aplicações efetivas na construção *Web Semântica* e nos ambientes colaborativos de inteligência coletiva.



CAPÍTULO 3



**MARCONT: UM DIÁLOGO
POSSÍVEL COM MARC21?**

3 MARCONT: um diálogo possível com MARC 21?

O que é bem conhecido, justamente por ser bem conhecido, não é conhecido. *Heigel*

Este capítulo apresenta o estudo minucioso e aprofundado da iniciativa MarcOnt, quais as principais características e sua identificação na sua essência, bem como a sua repercussão no cenário atual, no que se conhece hoje, como *Web Semântica*, uma proposta de Berners-Lee (2005), na tentativa de comparar os principais conceitos dos recursos (classes) contemplados pela iniciativa com os instrumentos de descrição do domínio bibliográfico.

A tese tem como propostas e objetivos iniciais o estudo e a análise do MarcOnt e suas interfaces em relação ao padrão de metadados MARC 21, de domínio da Biblioteconomia e Ciência da Informação. Nesse sentido, verifica-se se tal comparação é possível.

Para análise desse estudo, faz-se necessário retomar primeiramente como foi pensada a iniciativa MarcOnt, seu contexto, momento histórico etc. Dessa forma, procura-se abordar e recrutar elementos para o desenvolvimento do estudo do objeto de investigação dessa tese.

3.1 Gênese da iniciativa MarcOnt

Aumentar o investimento em pesquisa e desenvolvimento, bem como a tendência de produzir informação de qualidade contemporaneamente, resultou em um mais rápido crescimento exponencial de recursos disponíveis na rede. Informações de alta qualidade estão armazenadas, representadas e preservadas em bibliotecas digitais, que cada vez mais, se dedicam e empenham seus esforços, no sentido de expandir “ilhas” de informação bem estruturadas, organizadas e descritas. Para proporcionar a interoperabilidade entre bibliotecas, ambientes informacionais e sistemas de informação, uma alternativa é o investimento no desenvolvimento de padrões de descrição bibliográfica, que garantirão a representação e a padronização dos recursos informacionais.

Nesse cenário, e de acordo com os autores Kruk, Synak e Zimmermann (2005a), a iniciativa MarcOnt foi pensada num primeiro momento, com o objetivo de desenvolver uma ontologia para a descrição bibliográfica e relacionar as ferramentas e os instrumentos relacionados às tecnologias da *Web Semântica*, baseada no legado de três padrões de descrição bibliográfica: MARC 21 de domínio da Biblioteconomia, Dublin Core para localização de recursos na *Web* e BibTeX para a descrição de recursos na Internet e aplicado em bibliotecas digitais atualmente.

Nesse momento vale frisar que a iniciativa MarcOnt foi construída do ponto de vista da Ciência da Computação, onde seu idealizador, Sebastian Kruk, numa revisão de literatura procurou, através do estado da arte no domínio biblioteconômico, identificar qual ou quais eram os padrões mais utilizados nessa área e se o mesmo era compatível com o MarcOnt, e identificou, de acordo com seu estudo que o padrão MARC 21 era o mais aceito internacionalmente pelos bibliotecários.

As novas soluções não podem negligenciar padrões e formatos existentes ou nunca se tornarão amplamente conhecidos. No momento, o formato de dados bibliográficos MARC21 é ainda o principal padrão utilizado em todas as bibliotecas do mundo. A ontologia que aspira ser seu sucessor (ou ao menos contrapartes semânticas) tem que oferecer não menos capacidade na descrição. (SYNAK, 2005, p. 63, tradução nossa).

Kruk, Synak e Zimmermann (2005) afirmam que alguns padrões, como o formato MARC 21, geralmente são aplicados e utilizados em bibliotecas clássicas (convencionais), enquanto as atuais bibliotecas digitais, que se configuram hoje, tendem a apoiar formatos semanticamente mais ricos, tais como Dublin Core ou BibTeX.

Contrapondo Kruk, Synak e Zimmermann (2005b) afirma-se o desconhecimento dos autores no que concerne ao padrão de metadados MARC 21, pois este é considerado pela classificação dos padrões de metadados (DEMPSEY; HEERY, 1997) um formato rico, apresentando um nível de detalhamento na descrição de forma completa e complexa, baseado na semântica do AACR2 (Código de Catalogação Anglo Americano) e exige a necessidade de um profissional (catalogador/especialista) na confecção de representações que contemple a integridade das informações. Enquanto os padrões de metadados Dublin Core e BibTeX podem ser considerados formatos estruturados e proporcionam uma descrição mínima do

recurso informacional, não expressando semânticas ricas, e ainda ser elaborado por profissionais não especialistas.

As bibliotecas digitais que se conhecem hoje, encontradas na Internet refletem esse cenário, por estarem representadas em formatos estruturados. Muitas vezes os sistemas de bibliotecas digitais suportam apenas um dos formatos de descrição que é determinado para um público específico, como mencionam Kruk, Synak e Zimmermann (2005c, p. 2, tradução nossa), ao classificarem as bibliotecas digitais em algumas tipologias, tais como:

Bibliotecas digitais clássicas (convencionais): os sistemas de bibliotecas para manipulação de recursos físicos armazenados em bibliotecas clássicas (convencionais), muitas vezes fornece interfaces *Web* adicionais, onde os leitores podem pesquisar e reservar livros selecionados. Em muitos casos MARC 21 é o único formato de descrição bibliográfica utilizado para a comunicação.

Bibliotecas digitais editoriais: fornece o acesso às publicações, conferências, *proceedings* etc. Uma vez que estes recursos são utilizados pelos pesquisadores, principalmente a maioria das bibliotecas digitais desse tipo apoiam o formato de descrição BibTeX.

Outras bibliotecas digitais: fornecem diferentes tipos de recursos direcionados para um tipo diferente de usuário ou apenas usuários da Internet. Muitas dessas aplicações *Web* utilizam o padrão de metadados Dublin Core para anotar o recurso apresentado.

Bibliotecas digitais semânticas: com o alvorecer da *Web Semântica* as bibliotecas digitais estão mais habilitadas semanticamente para emergir nesse contexto, como a JeromeDL. O objetivo principal dessas bibliotecas digitais é proporcionar melhores características de recuperação na interação homem-máquina e uma maior precisão na busca distribuída em redes de bibliotecas digitais heterogêneas.

As bibliotecas digitais se apresentam e atuam como um segmento na Internet onde procuram desenvolver e criar métodos e técnicas para a padronização dos recursos informacionais. Mas garantir/assegurar a interoperabilidade entre “ilhas” de informações, bem estruturadas e padronizadas, entre formatos de descrição bibliográfica distintos encontra-se numa questão de investigação e preocupação pela comunidade científica.

Novas tecnologias baseadas em pesquisas na temática *Web Semântica* desenvolvem-se com o intuito de tentar agregar semântica à descrição bibliográfica para um tratamento mais efetivo dos conteúdos digitais. A *Web Semântica* oferece algumas soluções para diferentes problemas criando uma nova visão para o armazenamento e o processamento dos dados.

Algumas dessas soluções poderiam ser implementadas para potencializar os resultados das buscas no âmbito das bibliotecas digitais.

Quando se pensa em agregar semântica à descrição dos recursos em bibliotecas digitais e sua compatibilidade para que ocorra a interoperabilidade, faz-se necessário o desenvolvimento de uma ontologia comum. A ontologia pode ser considerada como um instrumento para a descrição bibliográfica dos recursos informacionais, pois se cria uma rede de conceitos com propriedades apropriadas e restritas. Kruk, Synak e Zimmerman (2005b, p. 1, tradução nossa) alegam que “infelizmente, é muito difícil convencer os bibliotecários para realizarem o esforço da criação da descrição semântica dos recursos, como por exemplo, os livros”.

Vale destacar que os autores referem-se a esta afirmação em 2005, uma vez que atualmente (2012), muitos bibliotecários, principalmente da *Library of Congress* e dos organismos internacionais de catalogação (IFLA), têm trabalho em parceria com profissionais da Ciência da Computação, a fim de desenvolverem técnicas e ferramentas que viabilizem a explicitação da semântica nos relacionamentos bibliográficos, que será tratado nos capítulos 4 e 5.

Como aponta Synak (2005), muitas são as razões para a construção de um novo padrão para a descrição bibliográfica semântica, dentre as quais podemos citar o argumento de que o padrão de metadados MARC 21 não fornece informação semântica necessária de forma a ser compreendida por uma máquina.

Nesse contexto, inúmeros trabalhos identificam problemas de semântica nos formatos de descrição bibliográfica, de modo que a utilização de ontologias em ambientes informacionais digitais, tais como as bibliotecas digitais vêm das necessidades reais em vez da investida comum atualmente de usar uma tecnologia, simplesmente porque ela existe. (DABROWSKI; SYNAK; KRUK, 2009).

O ponto principal é que se bem construída uma ontologia pode representar uma visão de mundo, potencializando representações semânticas que não poderiam ser obtidas a partir de descrições textuais. As representações semânticas podem fornecer informações sobre os objetos na maneira como as pessoas tendem a pensar. Nós usamos associações livres,

agrupamento de objetos por semelhanças e permitindo muitos tipos de contextos, diferenciando a importância dos fatores de acordo com a situação. Uma foto aérea vai significar algo diferente para um turista do que para um militar. (DABROWSKI; SYNAK; KRUK, 2009).

Ressalta-se pelos apontamentos dos autores a analogia que pode ser feita em relação às formas de representação da informação, que pode ser visualizada no perspectivismo de Peterson (1996), nas construções mentais dos sujeitos psicossociais na realização de uma tarefa específica.

Na perspectiva de Peterson (1996) formas de representação podem ser vistas como um tema interdisciplinar à ciência cognitiva, pois há uma grande multiplicidade de formas de representação (anotações, formalismos, interfaces, linguagens de programação etc.), e a admissão de uma nova forma de representação pode tornar um domínio mais acessível para um determinado propósito, comunidades de interesse específico ou tipo de usuário.

De acordo com Peterson (1996), a adequação de uma perspectiva é sempre relativa, mas ela não se estabelece de modo arbitrário. Na escolha de uma perspectiva deve-se considerar a sua adequação ao usuário a uma situação determinada, a um processo, a uma ontologia e a uma meta específica. É a análise de cada um destes componentes que se pretende desenvolver no estudo dos processos que atuam nas diretrizes, modelagens e estruturas de sistemas para atendimento de necessidades de sujeitos em ambientes informacionais específicos.

Nesse sentido, avaliar a teoria das formas de representação como um fundamento teórico-metodológico para analisar e definir uma forma de representação bibliográfica e documentária, em face de uma determinada tarefa se justifica na avaliação das aplicações práticas dos padrões de representações catalográficas e de modo especial, de formatos de intercâmbio de dados bibliográficos e catalográficos e dos padrões de metadados que garantem a interoperabilidade.

Diferentes formas de representação podem ainda ser usadas para apresentar um objeto ou conceito de diferentes pontos de vista. (CASA, 1997). O aspecto positivo está justamente na possibilidade de obter “estruturas capazes de representar somente a informação relevante

em um determinado momento”. (CASA, 1997, p. 210). Podemos entender como estrutura neste contexto, algo que propicie e dê suporte informacional.

Pode-se dizer que a representação da informação registrada é compreendida numa perspectiva de que é primeiramente mentalizada ou internalizada numa estrutura cerebral, onde os sujeitos cognoscentes (mediadores ou intermediários), precisam externalizar em um formato ou em um suporte informacional, que garanta o registro (armazenamento) do conhecimento para uma posterior socialização, uso e (re)uso dos recursos informacionais. (CASTRO, 2008).

A *Web Semântica* compreende a idéia de construir e compartilhar conhecimento, ou seja, vários tipos e formas de representação de um recurso informacional. Ontologias são construídas para fornecer a descrição de um dado domínio de interesse. Depois de atingido um acordo comum sobre uma descrição, ela torna-se um padrão (a ontologia FOAF²⁵ é um bom exemplo). A construção de ontologias deve reutilizar os conhecimentos adequados de outras ontologias já existentes, se possível. Esta abordagem é especialmente importante no domínio das bibliotecas digitais, onde múltiplos padrões de metadados são formulados e amplamente utilizados, mencionam-se os padrões MARC 21, Dublin Core e BibTeX. Tendo em mente as metas de uma ontologia bibliográfica, os autores decidem construir primeiro um projeto, com base em regras, normas e padrões existentes para a descrição de recursos bibliográficos. (DABROWSKI; SYNAK; KRUK, 2009).

A construção de uma ontologia bibliográfica que garanta ser “fiel” às normas, aos padrões e aos hábitos de comunidades específicas proporcionando capacidades superiores de tecnologias semânticas é o primeiro passo para assegurar a interoperabilidade semântica entre os ambientes informacionais digitais. O que implicará mudanças nas formas de pesquisa, de acesso e de utilização de recursos que se conhecem hoje. (DABROWSKI; SYNAK; KRUK, 2009).

²⁵ FOAF (*Friend of a Friend*) permite a descrição de pessoas e organizações/instituições responsáveis pela criação do recurso. No caso de uma ontologia ela diz respeito ao processo de criação ou publicação de um recurso bibliográfico. Maiores detalhes podem ser encontrados na documentação técnica do projeto FOAF. Disponível em: <<http://www.foaf-project.org/original-intro>>. Acesso em: 21 jan. 2012.

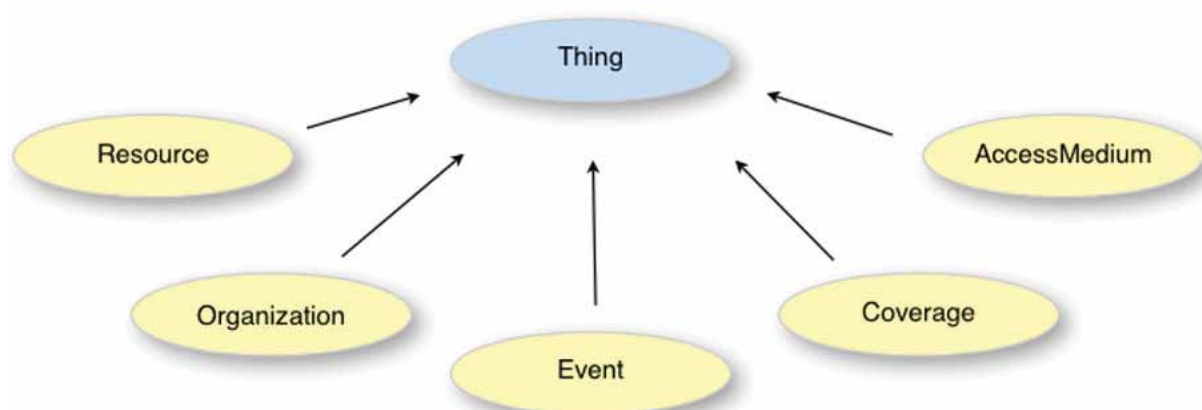
Para uma compreensão mais detalhada da ontologia MarcOnt faz-se necessário dizer que esta possui uma estrutura definida em classes e propriedades utilizadas para a classificação de recursos bibliográficos.

3.2 Estrutura do MarcOnt: classes e propriedades

É oportuno e relevante destacar novamente que a principal meta da ontologia MarcOnt é prover um formato de descrição bibliográfica uniforme, capturando conceitos de formatos já existentes, tais como BibTeX, Dublin Core e MARC 21.

Quando falamos em ontologias, as classes são frequentemente identificadas com objetos do mundo real. Seus nomes muitas vezes refletem este caso (por exemplo, a classe Pessoa na ontologia FOAF representa um ser humano). A criação de listas de classes na ontologia e sua hierarquia parecem ser simples. O principal problema ocorre quando se deve construir um modelo de um domínio de interesse particular, em vários modelos existentes neste domínio. Tal ontologia deve ser construída com base na existência de padrões de metadados existentes, tais como MARC21, Dublin Core e BibTeX, o que implica ou gera um processo complicado de obtenção de consenso. (DABROWSKI; SYNAK; KRUK, 2009). Nesse sentido foram definidas as seguintes classes na ontologia MarcOnt, de acordo com a figura 7.

FIGURA 7: Estrutura geral de classes da ontologia MarcOnt.



Fonte: Dabrowski, Synak e Kruk (2009, p. 116).

✚ **Resource:** usado como base para representar todos os tipos de materiais que podem ser armazenados em uma biblioteca. É considerada a classe mais importante na ontologia MarcOnt. As especificações de BibTeX tiveram grande impacto sobre a hierarquia dos recursos, pois de acordo com Dabrowski, Synak e Kruk (2009) é um formato amplamente usado para descrever recursos bibliográficos onde os autores dos recursos bibliográficos poderiam fornecer suporte ao BibTeX.

✚ **Organization:** usado para representar a filiação de um agente (foaf:Agent) ou para representar a estrutura organizacional de um dado recurso digital. Com base no caso de uso da *Digital Enterprise Research Institute* (DERI²⁶) identificou-se os conceitos de modelagem hierárquica organizacional em DERI.

✚ **Event:** esta classe representa todos os eventos importantes da perspectiva do processo de publicação e também eventos onde os determinados trabalhos foram publicados ou apresentados. Existem quatro tipos principais de eventos de acordo com a ontologia MarcOnt:

- *Conferência:* representa o caso de conferências muitas vezes usado para definir o local de publicação onde um dado trabalho foi apresentado.
- *Workshop:* modelos conceitos de workshop onde determinando trabalho foi apresentado.
- *Apresentação:* caso de apresentação de uma determinada obra ou trabalho. Identificou-se uma série de diferentes tipos de apresentação, incluindo: Tutorial, Sessão oral e Sessão de pôsteres.
- *Reunião/Encontro:* representa o conceito de uma reunião ou encontro formal ou informal (exemplo, uma reunião onde o conteúdo de um relatório técnico foi aprovado).

✚ **Coverage:** fornece informações sobre a cobertura internacional de um determinado recurso informacional (ex.: Internacional, Nacional etc.).

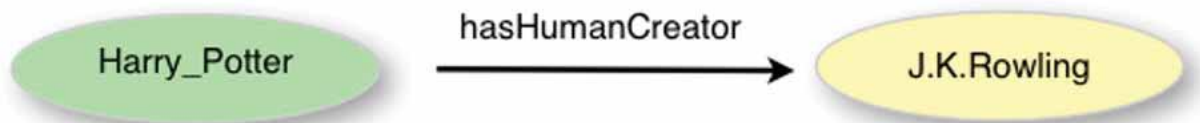
²⁶ Mais especificações sobre *Digital Enterprise Research Institute* (DERI) podem ser encontradas no *site*: Disponível em: < <http://www.deri.ie/>>. Acesso em: 21 maio de 2010.

- ✚ *AccessMedium*: representa diferentes tipos de mídias disponíveis para acessar um recurso informacional (ex.: o artigo de um jornal específico pode ser acessado como uma impressão, ou como uma versão eletrônica).

Enquanto as classes representam tipos de objetos em um determinado domínio específico, as classes podem ser entendidas como seus atributos. Destaca-se que no domínio bibliográfico os atributos podem ser entendidos como as características de um objeto. São utilizados em um conjunto de elementos de metadados e definem as propriedades dos elementos.

O exemplo simples de um dos autores do livro pode ser compreendido de acordo com a figura 8.

FIGURA 8: Recurso informacional livro com seus atributos.



Fonte: Dabrowski, Synak e Kruk (2009, p. 116).

A figura 8 mostra um exemplo clássico de um livro, ou seja, um objeto (um livro cujo título é Harry Potter) que possui um atributo (propriedade *hasHumanCreator*), ligando-o ao objeto autor (J. K. Rowling) que representa a classe autor do recurso bibliográfico.

Além das classes, durante o desenvolvimento da ontologia MarcOnt foram construídas um conjunto maior de propriedades, que podem ser resumidas de acordo com o Quadro 5.

Quadro 5: Principais propriedades da ontologia MarcOnt.

Propriedades	Definição
hasPublisher	Usado para descrever um publicador ou editor de um determinado recurso.
hasCreator	Esta propriedade pode ser usada para descrever o criador de um determinado recurso ou de determinadas coleções de recursos. Pode ser uma pessoa, um grupo ou uma organização.
hasHumanCreator	esta propriedade pode ser usada para descrever o criador humano (uma pessoa) de um determinado recurso ou de determinadas coleções de recursos.
hasCopyright	Esta propriedade confere os direitos autorais de um determinado recurso.
hasSource	Usado como uma propriedade citação para descrever uma fonte de referência. Permite criar referências entre recursos.
hasDomain	Descreve o domínio de interesse apropriado para o recurso.
hasTagging	Refere-se ao conceito de marcação na ontologia ttm e é baseado em TagCommons. Permite fazer referência a uma marcação realizada por um usuário anexando tags ao recurso.

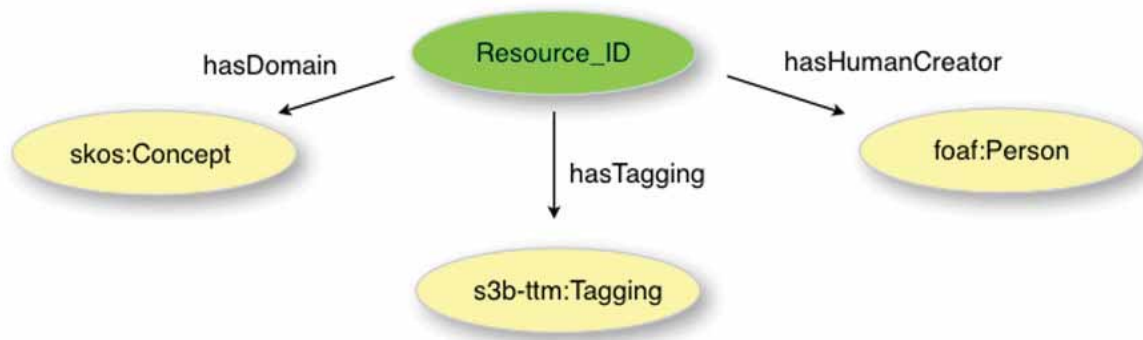
Fonte: Elaborado pelo autor.

Muitas outras propriedades foram definidas na ontologia, mas foram apontadas as principais num estudo inicial. Destaca-se que as principais classes e propriedades da ontologia MarcOnt foi orientada para suportar os padrões de metadados Dublin Core, BibTeX e MARC 21, utilizados também na Biblioteca Digital Semântica JeromeDL²⁷.

A figura 9 apresenta a descrição de um recurso bibliográfico utilizando a ontologia MarcOnt.

²⁷ A iniciativa MarcOnt nasceu dos primeiros experimentos na Biblioteca Digital JeromeDL. Detalhes específicos podem ser encontrados no portal. Disponível em: <<http://www.jeromedl.org/>>. Acesso em: 23 abr. 2010.

FIGURA 9: Descrição de um recurso bibliográfico utilizando a ontologia MarcOnt.



Fonte: Dabrowski, Synak e Kruk (2009, p. 119).

A classe Recurso (*Resource_ID*) pode ser identificada como o núcleo da ontologia MarcOnt. Um recurso bibliográfico particular foi desenvolvido pelo seu criador. A propriedade *hasHumanCreator* é usada para conectar uma pessoa (representada com a classe *foaf:Person*) a um determinado recurso. A criação de um recurso muitas vezes está relacionada à sua propriedade, no sentido de um conjunto de direitos exclusivos regulamentam a utilização de um recurso em especial – *copyright*. As propriedades *hasDomain* e *hasTagging* estão disponíveis para fornecer anotações semânticas de um recurso. Pode-se usar a propriedade *hasTagging* para anexar uma marcação à um recurso. *Tagging* é identificada com a ontologia S3B²⁸ cobrindo todos os aspectos do processo de marcação (criador de uma *tag*, *tags* utilizadas, momento em que ocorreu a marcação). Por outro lado, a classificação de um recurso é possível com o uso da propriedade *hasDomain*. Esta propriedade representa um domínio de interesse de um recurso e permite uma ligação com outro recurso através de *skos:Concept*²⁹.

Em um índice alfabético estão os principais termos da ontologia MarcOnt divididos por classes (categorias ou tipos) e por propriedades dos recursos informacionais, apresentados por Dabrowski e Kruk (2007) nas especificações iniciais do projeto, listadas a seguir.

²⁸ A documentação de *S3B Tagging Ontology* pode ser encontrada no site: Disponível em: <<http://s3b.corrib.org/tagging/spec/>>. Acesso em: 30 maio 2010.

²⁹ *Simple Knowledge Organization System* (SKOS) é um modelo de dados para o compartilhamento e a ligação de sistemas de organização do conhecimento através da *Web* Semântica.

CLASSES

<p>1. Class: <code>marcont:#Article</code></p>	<p><i>Article</i> – um dos tipos de recursos (artigo) É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td>sub-class-of:</td> <td><code>marcont:Resource</code></td> </tr> <tr> <td>in-domain-of:</td> <td><code>marcont:hasJournal</code></td> </tr> </table> <p>No que diz respeito à classe <i>article</i>, nota-se que esta possui uma subclasse (pode ser considerada também uma classe) e que faz parte do domínio <i>Journal</i></p>	sub-class-of:	<code>marcont:Resource</code>	in-domain-of:	<code>marcont:hasJournal</code>
sub-class-of:	<code>marcont:Resource</code>				
in-domain-of:	<code>marcont:hasJournal</code>				
<p>2. Class: <code>marcont:#Book</code></p>	<p><i>Books</i> - um dos tipos de recursos (livros) É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td>sub-class-of:</td> <td><code>marcont:Resource</code></td> </tr> </table> <p>A classe <i>book</i> possui também uma subclasse <i>Resource</i>, ou seja, também pode ser considerado um recurso informacional.</p>	sub-class-of:	<code>marcont:Resource</code>		
sub-class-of:	<code>marcont:Resource</code>				
<p>3. Class: <code>marcont:#Booklet</code></p>	<p><i>Booklet</i> - um dos tipos de recursos É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td>sub-class-of:</td> <td><code>marcont:Resource</code></td> </tr> </table> <p>A classe <i>book</i> possui também uma subclasse <i>Resource</i>, ou seja, também pode ser considerado um recurso informacional.</p>	sub-class-of:	<code>marcont:Resource</code>		
sub-class-of:	<code>marcont:Resource</code>				
<p>4. Class: <code>marcont:#Chapter</code></p>	<p><i>Chapter</i> – esta classe representa um capítulo de livro. É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td>sub-class-of:</td> <td><code>marcont:StructureElement</code></td> </tr> <tr> <td>in-range-of:</td> <td><code>marcont:hasChapter</code></td> </tr> </table> <p>A classe <i>Chapter</i> possui uma subclasse de um elemento estrutural de um recurso bibliográfico.</p>	sub-class-of:	<code>marcont:StructureElement</code>	in-range-of:	<code>marcont:hasChapter</code>
sub-class-of:	<code>marcont:StructureElement</code>				
in-range-of:	<code>marcont:hasChapter</code>				

CLASSES

<p>5. Class: marcont:#Cluster</p>	<p><i>Cluster</i> – representa um conceito de <i>cluster</i> – subunidade de Instituto. É subdividida em:</p> <table border="1" data-bbox="762 461 1238 506"> <tr> <td>sub-class-of:</td> <td><u>marcont:Organization</u></td> </tr> </table> <p>A classe Cluster possui uma subclasse Organization, também pode ser considerado um recurso informacional.</p>	sub-class-of:	<u>marcont:Organization</u>				
sub-class-of:	<u>marcont:Organization</u>						
<p>6. Class: marcont:#Collection</p>	<p><i>Collection</i> – representa coleções de recursos. É subdividida em:</p> <table border="1" data-bbox="762 797 1238 842"> <tr> <td>sub-class-of:</td> <td><u>marcont:Resource</u></td> </tr> </table> <p>A classe <i>collection</i> possui também uma subclasse <i>Resource</i>, ou seja, também pode ser considerado um recurso informacional.</p>	sub-class-of:	<u>marcont:Resource</u>				
sub-class-of:	<u>marcont:Resource</u>						
<p>7. Class: marcont:#Conference</p>	<p><i>Conference</i> – um dos tipos de recursos. É subdividida em:</p> <table border="1" data-bbox="762 1169 1238 1214"> <tr> <td>sub-class-of:</td> <td><u>marcont:Resource</u></td> </tr> </table> <p>A classe <i>conference</i> possui também uma subclasse <i>Resource</i>, ou seja, também pode ser considerado um recurso informacional.</p>	sub-class-of:	<u>marcont:Resource</u>				
sub-class-of:	<u>marcont:Resource</u>						
<p>8. Class: marcont:#Copyright</p>	<p><i>Copyrights</i> – esta classe representa a entidade de direitos de autor. É subdividida em:</p> <table border="1" data-bbox="762 1527 1238 1671"> <tr> <td>in-range-of:</td> <td><u>marcont:hasCopyright</u></td> </tr> <tr> <td rowspan="3">in-domain-of:</td> <td><u>marcont:hasOwner</u></td> </tr> <tr> <td><u>marcont:hasEndDate</u></td> </tr> <tr> <td><u>marcont:hasBeginDate</u></td> </tr> </table> <p>A classe Copyright está na cobertura copyright e está no domínio do dono dos direitos autorais, as datas de início e fim de um recurso informacional.</p>	in-range-of:	<u>marcont:hasCopyright</u>	in-domain-of:	<u>marcont:hasOwner</u>	<u>marcont:hasEndDate</u>	<u>marcont:hasBeginDate</u>
in-range-of:	<u>marcont:hasCopyright</u>						
in-domain-of:	<u>marcont:hasOwner</u>						
	<u>marcont:hasEndDate</u>						
	<u>marcont:hasBeginDate</u>						

CLASSES

<p>9. Class: marcont:#Coverage</p>	<p><i>Coverage</i> – esta classe representa o alcance da “cobertura” de propriedade de anotação advindo de Dublin Core (DC). É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td style="padding: 2px;">in-range-of:</td> <td style="padding: 2px;">marcont:hasCoverage</td> </tr> </table> <p>A classe <i>coverage</i> também pode ser considerada um recurso informacional e está no alcance de marcont:hasCoverage.</p>	in-range-of:	marcont:hasCoverage
in-range-of:	marcont:hasCoverage		
<p>10. Class: marcont:#DCMIType</p>	<p><i>DCMI type</i> – representa a lista de Dublin Core Initiative (DCMI) <i>types</i> que podem ser em escala da propriedade sahType. É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td style="padding: 2px;">in-range-of:</td> <td style="padding: 2px;">marcont:hasDCMIType</td> </tr> </table>	in-range-of:	marcont:hasDCMIType
in-range-of:	marcont:hasDCMIType		
<p>11. Class: marcont:#Faculty</p>	<p><i>Faculty</i> – representa um conceito de faculdade – subunidade de universidade. É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td style="padding: 2px;">sub-class-of:</td> <td style="padding: 2px;">marcont:Organization</td> </tr> </table>	sub-class-of:	marcont:Organization
sub-class-of:	marcont:Organization		
<p>12. Class: marcont:#Inbook</p>	<p><i>Inbook</i> – um dos tipos de recursos – parte de um livro, bem como um capítulo, seção. É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td style="padding: 2px;">sub-class-of:</td> <td style="padding: 2px;">marcont:Resource</td> </tr> </table>	sub-class-of:	marcont:Resource
sub-class-of:	marcont:Resource		

CLASSES

<p>13. Class: marcont:#Incollection</p>	<p>Incollection – um dos tipos de recursos – parte de um livro com o seu próprio título. É subdividida em:</p> <table border="1" data-bbox="762 539 1238 584"> <tr> <td>sub-class-of:</td> <td><u>marcont:Resource</u></td> </tr> </table>	sub-class-of:	<u>marcont:Resource</u>				
sub-class-of:	<u>marcont:Resource</u>						
<p>14. Class: marcont:#Inproceedings</p>	<p><i>Inproceedings</i> – um dos tipos de recursos – um artigo publicado nos anais de evento. É subdividida em:</p> <table border="1" data-bbox="762 909 1238 954"> <tr> <td>sub-class-of:</td> <td><u>marcont:Resource</u></td> </tr> </table>	sub-class-of:	<u>marcont:Resource</u>				
sub-class-of:	<u>marcont:Resource</u>						
<p>15. Class: marcont:#Institute</p>	<p><i>Institute</i> – representa o conceito de instituto. É subdividida em:</p> <table border="1" data-bbox="762 1261 1238 1305"> <tr> <td>sub-class-of:</td> <td><u>marcont:Organization</u></td> </tr> </table>	sub-class-of:	<u>marcont:Organization</u>				
sub-class-of:	<u>marcont:Organization</u>						
<p>16. Class: marcont:#Journal</p>	<p><i>Journal</i> – representa a entidade revista. É subdividida em:</p> <table border="1" data-bbox="762 1592 1238 1715"> <tr> <td>sub-class-of:</td> <td><u>marcont:Collection</u></td> </tr> <tr> <td>in-range-of:</td> <td><u>marcont:hasJournal</u></td> </tr> <tr> <td>in-domain-of:</td> <td><u>marcont:hasNumber</u></td> </tr> </table>	sub-class-of:	<u>marcont:Collection</u>	in-range-of:	<u>marcont:hasJournal</u>	in-domain-of:	<u>marcont:hasNumber</u>
sub-class-of:	<u>marcont:Collection</u>						
in-range-of:	<u>marcont:hasJournal</u>						
in-domain-of:	<u>marcont:hasNumber</u>						

CLASSES

17. Class: marcont:#Lab	<p><i>Laboratory</i> – representa um conceito de laboratório – subunidade de <i>cluster</i>. É subdividida em:</p> <table border="1"><tr><td>sub-class-of:</td><td>marcont:Organization</td></tr></table>	sub-class-of:	marcont:Organization
sub-class-of:	marcont:Organization		
18. Class: marcont:#Manual	<p><i>Manual</i> – um dos tipos de recursos – documentação técnica. É subdividida em:</p> <table border="1"><tr><td>sub-class-of:</td><td>marcont:Resource</td></tr></table>	sub-class-of:	marcont:Resource
sub-class-of:	marcont:Resource		
19. Class: marcont:#Mastersthesis	<p><i>Mastersthesis</i> – um dos tipos de recursos. É subdividida em:</p> <table border="1"><tr><td>sub-class-of:</td><td>marcont:Resource</td></tr></table>	sub-class-of:	marcont:Resource
sub-class-of:	marcont:Resource		
20. Class: marcont:#Misc	<p><i>Misc</i> – um dos tipos de recursos. É subdividida em:</p> <table border="1"><tr><td>sub-class-of:</td><td>marcont:Resource</td></tr></table>	sub-class-of:	marcont:Resource
sub-class-of:	marcont:Resource		

CLASSES

<p>21. Class: marcont:#Month</p>	<p>esta classe representa o tipo de enumeração englobando todos os valores apropriados para meses. É subdividida em:</p> <table border="1" data-bbox="762 521 1238 566"> <tr> <td>in-range-of:</td> <td>marcont:hasMonth</td> </tr> </table>	in-range-of:	marcont:hasMonth
in-range-of:	marcont:hasMonth		
<p>22. Class: marcont:#Organization</p>	<p><i>Organization</i> – representa o conceito de organização. É subdividida em:</p> <table border="1" data-bbox="762 860 1238 936"> <tr> <td>in-range-of:</td> <td>marcont:hasAffiliation marcont:hasOrganization</td> </tr> </table>	in-range-of:	marcont:hasAffiliation marcont:hasOrganization
in-range-of:	marcont:hasAffiliation marcont:hasOrganization		
<p>23. Class: marcont:#Phdthesis</p>	<p><i>Phdthesis</i> – um dos tipos de recursos. É subdividida em:</p> <table border="1" data-bbox="762 1240 1238 1285"> <tr> <td>sub-class-of:</td> <td>marcont:Resource</td> </tr> </table>	sub-class-of:	marcont:Resource
sub-class-of:	marcont:Resource		
<p>24. Class: marcont:#Proceedings</p>	<p><i>Proceedings</i> - um dos tipos de recursos. É subdividida em:</p> <table border="1" data-bbox="762 1612 1238 1657"> <tr> <td>sub-class-of:</td> <td>marcont:Collection</td> </tr> </table>	sub-class-of:	marcont:Collection
sub-class-of:	marcont:Collection		

CLASSES

<p>25. Class: marcont:#Project</p>	<p><i>Project</i> – representa o conceito de projeto É subdividida em:</p> <table border="1" style="width: 100%;"> <tr> <td style="width: 30%;">sub-class-of:</td> <td>marcont:Organization</td> </tr> </table>	sub-class-of:	marcont:Organization		
sub-class-of:	marcont:Organization				
<p>26. Class: marcont:#ReferableEntity</p>	<p><i>referable entity</i> – essa classe representa uma entidade que pode ser referenciada. É subdividida em:</p> <table border="1" style="width: 100%;"> <tr> <td style="width: 30%;">sub-class-of:</td> <td>marcont:StructureElement</td> </tr> <tr> <td>in-range-of:</td> <td>marcont:hasCrossRef</td> </tr> </table>	sub-class-of:	marcont:StructureElement	in-range-of:	marcont:hasCrossRef
sub-class-of:	marcont:StructureElement				
in-range-of:	marcont:hasCrossRef				
<p>27. Class: marcont:#Resource</p>	<p><i>Resource</i> – classe base para todos os recursos bibliográficos É subdividida em:</p> <table border="1" style="width: 100%;"> <tr> <td style="width: 30%;">in-range-of:</td> <td> marcont:hasSource marcont:hasPublisher marcont:hasContents marcont:hasHumanCreator marcont:hasCoverage marcont:hasChapter marcont:hasSource marcont:hasCopyright marcont:hasDomain marcont:hasDCMIType marcont:hasContributor marcont:hasTopic marcont:hasKeyword marcont:hasCrossRef marcont:hasOrganization </td> </tr> <tr> <td>in-domain-of:</td> <td> marcont:hasSponsor marcont:hasTagging marcont:hasSchool marcont:hasIdentifier marcont:hasOrder marcont:hasYear marcont:hasNote marcont:hasEdition marcont:hasISSN marcont:hasAddress marcont:hasAbstract marcont:howPublished marcont:hasISBN marcont:hasSeries </td> </tr> </table>	in-range-of:	marcont:hasSource marcont:hasPublisher marcont:hasContents marcont:hasHumanCreator marcont:hasCoverage marcont:hasChapter marcont:hasSource marcont:hasCopyright marcont:hasDomain marcont:hasDCMIType marcont:hasContributor marcont:hasTopic marcont:hasKeyword marcont:hasCrossRef marcont:hasOrganization	in-domain-of:	marcont:hasSponsor marcont:hasTagging marcont:hasSchool marcont:hasIdentifier marcont:hasOrder marcont:hasYear marcont:hasNote marcont:hasEdition marcont:hasISSN marcont:hasAddress marcont:hasAbstract marcont:howPublished marcont:hasISBN marcont:hasSeries
in-range-of:	marcont:hasSource marcont:hasPublisher marcont:hasContents marcont:hasHumanCreator marcont:hasCoverage marcont:hasChapter marcont:hasSource marcont:hasCopyright marcont:hasDomain marcont:hasDCMIType marcont:hasContributor marcont:hasTopic marcont:hasKeyword marcont:hasCrossRef marcont:hasOrganization				
in-domain-of:	marcont:hasSponsor marcont:hasTagging marcont:hasSchool marcont:hasIdentifier marcont:hasOrder marcont:hasYear marcont:hasNote marcont:hasEdition marcont:hasISSN marcont:hasAddress marcont:hasAbstract marcont:howPublished marcont:hasISBN marcont:hasSeries				

CLASSES

<p>28. Class: marcont:#StructureElement</p>	<p>Structure element – <i>esta classe representa o elemento estrutural de um recurso bibliográfico.</i></p>				
<p>29. Class: marcont:#TableOfContents</p>	<p><i>table of contents</i> – esta classe representa uma tabela de conteúdos É subdividida em:</p> <table border="1" data-bbox="762 869 1235 954"> <tr> <td>sub-class-of:</td> <td>marcont:StructureElement</td> </tr> <tr> <td>in-range-of:</td> <td>marcont:hasContents</td> </tr> </table>	sub-class-of:	marcont:StructureElement	in-range-of:	marcont:hasContents
sub-class-of:	marcont:StructureElement				
in-range-of:	marcont:hasContents				
<p>30. Class: marcont:#Techreport</p>	<p><i>Techreport</i> - um dos tipos de recursos – um relatório publicado por uma instituição ou universidade. É subdividida em:</p> <table border="1" data-bbox="762 1258 1235 1303"> <tr> <td>sub-class-of:</td> <td>marcont:Resource</td> </tr> </table>	sub-class-of:	marcont:Resource		
sub-class-of:	marcont:Resource				
<p>31. Class: marcont:#University</p>	<p><i>University</i> – esta classe representa o conceito de universidade É subdividida em:</p> <table border="1" data-bbox="762 1594 1235 1680"> <tr> <td>sub-class-of:</td> <td>marcont:Organization</td> </tr> <tr> <td>in-range-of:</td> <td>marcont:hasSchool</td> </tr> </table>	sub-class-of:	marcont:Organization	in-range-of:	marcont:hasSchool
sub-class-of:	marcont:Organization				
in-range-of:	marcont:hasSchool				

CLASSES			
32. Class: marcont:#Unpublished	<p><i>Unpublished</i> – um dos tipos de recursos – um documento nunca publicado, mas com um título e autor. É subdividida em:</p> <table border="1" style="margin-left: 20px;"> <tr> <td style="padding: 2px;">sub-class-of:</td> <td style="padding: 2px;">marcont:Resource</td> </tr> </table>	sub-class-of:	marcont:Resource
sub-class-of:	marcont:Resource		
33. Class: marcont:#hasMonth_Range	<p><i>has month range</i> – esta entidade representa a enumeração descrevendo os valores aceitos da propriedade hasMonth.</p>		
34. Class: marcont:#hasType_Range	<p>esta é um tipo de enumeração representando a cobertura da propriedade hasType.</p>		

Descritas as classes da ontologia MarcOnt, apresenta-se na seção seguinte suas propriedades com suas especificações.

PROPRIEDADES

1. Property: marcont:#hasAbstract - has abstract – resumo de um recurso.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

Destaca-se que toda vez que uma propriedade tiver cobertura (*Range*) *xsd:string*³⁰, encaminhará para *XML Schema* dada pelas recomendações do W3C.

2. Property: marcont:#hasAddress - has address – normalmente o endereço da casa publicadora ou outro tipo de instituição. Para grandes editoras, recomenda-se a omissão da informação completamente. Para pequenas editoras, por outro lado, pode-se ajudar o leitor, dando o endereço completo.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

3. Property: marcont:#hasAffiliation - has affiliation – afiliação dada a uma pessoa.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	foaf:Person
Range:	marcont:Organization

Destaca-se que toda vez que uma propriedade pertencer a um domínio do tipo *foaf:Person*³¹, significa que encaminhará para o namespace que possui as especificações de vocabulário FOAF.

4. Property: marcont:#hasAuthor - has Author – esta propriedade permite anexar informação sobre o autor de um recurso ou as coleções dos recursos.

É subdividida em:

OWL Type:	ObjectProperty
sub-property-of:	marcont:hasHumanCreator

³⁰ Disponível em: <<http://www.w3.org/2001/XMLSchema#related.resources>>. Acesso em: 12 maio 2010.

³¹ Disponível em: <<http://xmlns.com/foaf/spec/>>. Acesso em: 12 maio 2010.

PROPRIEDADES

5. Property: marcont:#hasBeginDate - has begin date – data de início do período de copyright.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Copyright
Range:	xsd:date

6. Property: marcont:#hasChapter – has chapter – descreve capítulos do recurso (livro, artigo etc.).

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:Chapter

7. Property: marcont:#hasContents - has table of contents – usado para identificar tabelas de conteúdos.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:TableOfContents

8. Property: marcont:#hasContributor - has contributor – uma entidade responsável por fazer contribuições ao conteúdo do recurso, mas não do autor.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	foaf:Agent

PROPRIEDADES

9. Property: marcont:#hasCopyright - has copyright – esta propriedade confere direitos autorais a um dado recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:Copyright

10. Property: marcont:#hasCoverage - has coverage – a extensão ou escopo do conteúdo do recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:Coverage

11. Property: marcont:#hasCreator - has creator – propriedade usada para descrever o criador do recurso ou a coleção dos recursos.

É subdividida em:

OWL Type:	ObjectProperty
Range:	foaf:Agent

12. Property: marcont:#hasCrossRef - has cross reference – equivalente à propriedade bibtex:hasCrossRef.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:ReferableEntity

PROPRIEDADES

13. Property: [marcont:#hasDCMIType](#) – has DCMI type – descreve tipos apropriados para o recurso de acordo com a iniciativa dicionário de tipos do Dublin Core (*Metadata Initiative dictionary of types*).

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:DCMIType

14. Property: [marcont:#hasDomain](#) – has domain of interest – descreve o domínio de interesse apropriado para o recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	skos:Concept

Destaca-se que toda propriedade que tiver como cobertura (Range) skos:Concept, encaminhará para as especificações do W3C SKOS³² *Simple Knowledge Organization System Namespace Document - HTML Variant*.

15. Property: [marcont:#hasEdition](#) - has edition – a edição de um livro, por exemplo, “Segunda”. Esta deve ser em ordinal, e deve ter a primeira letra em maiúscula, como mostrado aqui; o padrão styles converte em alguns casos para minúscula, quando necessário.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

16. Property: [marcont:#hasEditor](#) - has Editor – define o editor de um dado recurso ou coleção de recursos.

É subdividida em:

OWL Type:	ObjectProperty
sub-property-of:	marcont:hasHumanCreator

³² Disponível em: <<http://www.w3.org/TR/skos-reference/skos.html#semantic-relations>>. Acesso em: 12 maio 2010.

PROPRIEDADES

17. Property: marcont:#hasEndDate – has end date – a data de término do período de copyright.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Copyright
Range:	xsd:date

18. Property: marcont:#hasHumanCreator - has human creator – propriedade usada para descrever o criador do recurso ou coleção de recursos.

É subdividida em:

OWL Type:	ObjectProperty
sub-property-of:	marcont:hasCreator
Domain:	marcont:Resource
Range:	foaf:Person

19. Property: marcont:#hasISBN – has ISBN – esta propriedade permite atribuir um identificador na forma de ISBN.

É subdividida em:

OWL Type:	DatatypeProperty
sub-property-of:	marcont:hasIdentifier
Domain:	marcont:Resource

20. Property: marcont:#hasISSN – has ISSN – representa o número do ISSN atribuído a um dado recurso.

É subdividida em:

OWL Type:	DatatypeProperty
sub-property-of:	marcont:hasIdentifier
Domain:	marcont:Resource

PROPRIEDADES

21. Property: `marcont:#hasIdentifier` – has identifier – representa a identificação da propriedade.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

22. Property: `marcont:#hasJournal` - has journal – um nome do periódico. Abreviações estão previstas para muitas revistas; consulte o guia local.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Article
Range:	marcont:Journal

23. Property: `marcont:#hasKeyword` – has keyword – palavras-chave relacionadas ao recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	skos:Concept

24. Property: `marcont:#hasMonth` – has label – o mês em que a obra foi publicada, ou para um trabalho inédito, em que foi escrito. Deve-se utilizar a abreviatura padrão de três letras.

É subdividida em:

OWL Type:	ObjectProperty
Range:	marcont:Month

PROPRIEDADES

25. Property: [marcont:#hasNote](#) – *has note* – informação adicional que pode ajudar o leitor.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

26. Property: [marcont:#hasNumber](#) – *has number* – o número de uma revista, jornal, relatório técnico, ou de uma obra em uma série. Uma edição de um jornal ou revista é normalmente identificada pelo seu volume e número; a organização que elabora um relatório técnico geralmente atribui um número; em alguns livros são dados números em uma série nomeada.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Journal
Range:	xsd:int

27. Property: [marcont:#hasOrder](#) – *has order* – refere-se à ordem dos recursos na biblioteca digital JeromeDL. Cada recurso é atribuído um número para permitir o controle sobre sua ordem de armazenamento e descrição.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:int

28. Property: [marcont:#hasOrganization](#) – *has organization* – a organização que está envolvida num evento.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:Organization

PROPRIEDADES

29. Property: marcont:#hasOwner – has Owner – representa o proprietário dos direitos autorais.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Copyright
Range:	foaf:Agent

30. Property: marcont:#hasPages – has pages – número de páginas de um recurso ou capítulo.

É subdividida em:

OWL Type:	DatatypeProperty
Range:	xsd:int

31. Property: marcont:#hasPublisher – has publisher – descreve o editor de um determinado recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	foaf:Agent

32. Property: marcont:#hasSchool – has School – universidade onde o trabalho foi criado/publicado.

É subdividida em:

OWL Type:	ObjectProperty
sub-property-of:	marcont:hasOrganization
Domain:	marcont:Resource
Range:	marcont:University

PROPRIEDADES

33. Property: `marcont:#hasSeries` – has series – o nome de uma série ou conjunto de livros. Ao citar um livro no todo, o campo de título fornece o título e um campo opcional de série dá o nome de uma série ou conjuntos (vários volumes) em que o livro é publicado.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource

34. Property: `marcont:#hasSource` – has source – equivalente à propriedade *dc:source*.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	marcont:Resource

35. Property: `marcont:#hasSponsor` – has sponsor – patrocinador do recurso ou do evento.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	foaf:Agent

36. Property: `marcont:#hasTagging` – has tagging – tagging associada com o recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	ttm:Tagging

PROPRIEDADES

37. Property: marcont:#hasTitle – has title – o título da obra.

É subdividida em:

OWL Type:	DatatypeProperty
Range:	xsd:string

38. Property: marcont:#hasTopic – has topic – tópico de um recurso.

É subdividida em:

OWL Type:	ObjectProperty
Domain:	marcont:Resource
Range:	skos:Concept

39. Property: marcont:#hasURL – has URL – URI de um tipo DCMI de um recurso.

É subdividida em:

OWL Type:	DatatypeProperty
Range:	xsd:string

40. Property: marcont:#hasVolume – has volume – o volume de um jornal, revista ou de um livro em vários volumes.

É subdividida em:

OWL Type:	DatatypeProperty
Range:	xsd:string

PROPRIEDADES

41. Property: marcont:#hasYear – has year – o ano de publicação, ou para um trabalho inédito, o ano em que foi escrito. Geralmente, deve ser constituído de quatro algarismos. Ex.: 1984.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

42. Property: marcont:#howPublished – how published – como a obra foi publicada. A primeira palavra deve ser maiúscula.

É subdividida em:

OWL Type:	DatatypeProperty
Domain:	marcont:Resource
Range:	xsd:string

43. Property: marcont:#isPartOf – is part of – representa uma parte da relação entre os elementos que pertencem a uma coleção ou entre coleções.

É subdividida em:

OWL Type:	ObjectProperty
------------------	----------------

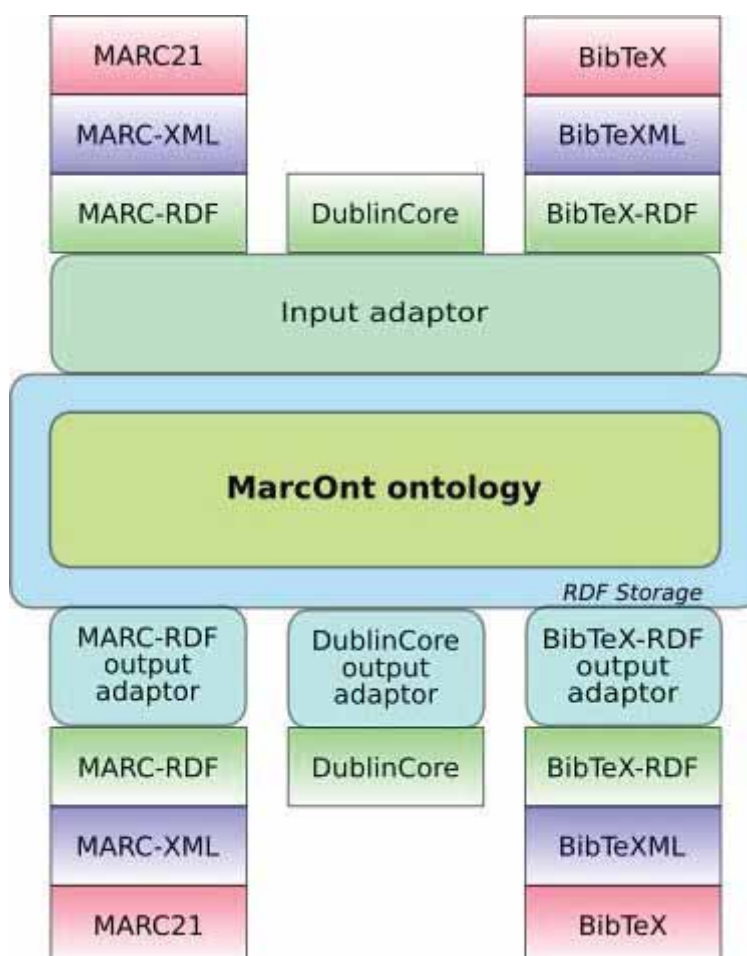
Identificadas e explicitadas as classes e as propriedades da ontologia MarcOnt, apresenta-se a arquitetura de mediação de serviços, onde seus idealizadores pensaram sobre as questões de interoperabilidade entre padrões de metadados heterogêneos.

3.3 Interoperabilidade entre padrões de metadados: o delineamento da arquitetura MarcOnt

Kruk, Synak e Zimmermann (2005a), para uma melhor compreensão da ontologia MarcOnt, propuseram uma arquitetura inicial a fim de ter uma visualização de como seria o seu funcionamento ao se pensar na questão da interoperabilidade semântica entre padrões de metadados heterogêneos, bem como o acesso aos recursos informacionais numa única interface.

A figura 10, em seguida, apresenta o delineamento de uma arquitetura inicial dos serviços de mediação MarcOnt.

FIGURA 10: Arquitetura de serviços de mediação MarcOnt.



Fonte: Kruk, Synak e Zimmermann (2005a, p. 2)

A arquitetura forma um fluxo, onde a parte superior é a entrada (*input*), as descrições semânticas estão no centro (*MarcOnt ontology*) e a parte mais inferior é a saída (*output*). Nota-se que o centro da arquitetura é o RDF (*Resource Description Framework*) – considerado o núcleo de armazenamento para descrições semânticas compiladas de outros padrões de metadados.

As caixas na primeira fileira (na entrada) e na última fileira (na saída) da arquitetura representam os formatos inerentes aos tradicionais (arquivos neste formato). As caixas no meio (MARCXML, BibTeXML) representam a primeira fase do processo de tradução, onde os arquivos inerentes são analisados gramaticalmente (sintaxe) e seus conteúdos são incluídos no formato XML.

Os adaptadores de entrada e saída (*input* e *output*) são implementados usando o mecanismo de inferência **Sesame**³³. De um lado dos adaptadores, temos as descrições semânticas e do outro, dados de descrições bibliográficas convertidas para um modelo RDF. Converter diferentes formatos de descrição para o modelo RDF depende do formato original do arquivo. Registros em formato MARC21 são armazenados em arquivos binários; então a conversão exige primeiro que tais arquivos sejam traduzidos para o formato MARCXML e então, numa fase posterior, convertem-se os dados para o modelo RDF usando XSLT (*Extensible Stylesheet Language*). Em relação ao *MarcOnt*, esta linguagem tem um papel importante principalmente no que diz respeito às formas de apresentação dos padrões de metadados contemplados no âmbito digital.

Nesse aspecto, vale ressaltar que Kruk (2005; 2006; 2007) destaca que ainda não há representações oficiais de dados de MARC 21 em RDF, desenvolvido pela *Library of Congress* (LC), e que os pesquisadores do *MarcOnt* criaram seu próprio formato (MARCRDF). “Apoiar outros formatos é uma questão de desenvolver adaptadores e conversores apropriados. Atualmente, estamos desenvolvendo adaptadores para MARC21”. (SYNAK, 2005, p. 75, tradução nossa). Contemporaneamente, a LC tem adaptado a versão

³³ Sesame é uma ferramenta, ou seja, um conjunto de classes-base desenvolvido em Java com código aberto para o desenvolvimento de sistemas, com o objetivo de armazenar e consultar dados em RDF. Este sistema é totalmente extensível e configurável no que concerne aos mecanismos de inferência e armazenamento de dados em formato RDF. Maiores informações sobre o sistema Sesame podem ser encontrados no *site*: Disponível em: <<http://www.openrdf.org/>>. Acesso em: 10 fev. 2012.

MARCXML para a arquitetura de metadados RDF, o que significa para a ontologia MarcOnt um ponto de partida para a interoperabilidade entre esquemas de metadados.

A fase final é traduzir os dados RDF para estruturas compatíveis com a ontologia MarcOnt. Os estágios precedentes requerem escrever ferramentas simples para que as bibliotecas as usem e possam intercambiá-las com relativa facilidade. (SYNAK, 2005, tradução nossa). Assim, pode-se dizer que essa arquitetura inicial visa ao compartilhamento entre padrões de metadados distintos, podendo proporcionar a interoperabilidade semântica entre ambientes informacionais digitais, onde as informações semânticas estarão disponíveis e armazenadas em uma única interface, neste caso, no MarcOnt.

3.3.1 Análise da arquitetura MarcOnt à luz da Catalogação Descritiva e dos princípios de interoperabilidade semântica

Para se entender melhor o propósito da ontologia MarcOnt, verifica-se como ela se dá em camadas para ser compreendida na sua totalidade.

FIGURA 11: RDF como *core* para descrições bibliográficas semânticas



Fonte: Adaptado de Kruk, Synak e Zimmermann (2005a, p. 2)

Para que se efetive a tão idealizada interoperabilidade semântica no visionário projeto MarcOnt, faz-se necessário uma correspondência entre os padrões de metadados vislumbrados nessa iniciativa. Para isso, como pode ser observado na figura 11, o núcleo de funcionamento está na camada central da arquitetura denominada *RDF Storage*. Pelos convencionais métodos de mapeamento entre esquemas de metadados, conforme visto no capítulo 2, nem sempre é possível que se configure a integridade dos dados no processo de

transformação. Para isso, na iniciativa MarcOnt foi desenvolvido um conversor chamado *RDF Translator*.

O adaptador/conversor *RDF Translator* contem algumas funcionalidades básicas que são relativamente fáceis para o uso. Além disso, algumas características adicionais podem ser facilmente adicionadas, caso necessário. Um exemplo dessas características são as funções adicionais que podem lidar com um URI ou *literals*³⁴.

Outro aspecto altamente modificável como parte das aplicações é o raciocínio lógico sobre os dados na entrada. Na primeira implementação um simples raciocínio foi utilizado. É fácil substituir o raciocínio com alguma solução mais potente, isso também inclui a possibilidade de dar ao usuário uma ampla gama de escolhas em relação aos esquemas de metadados. (NOVÁCEK et al., 2007).

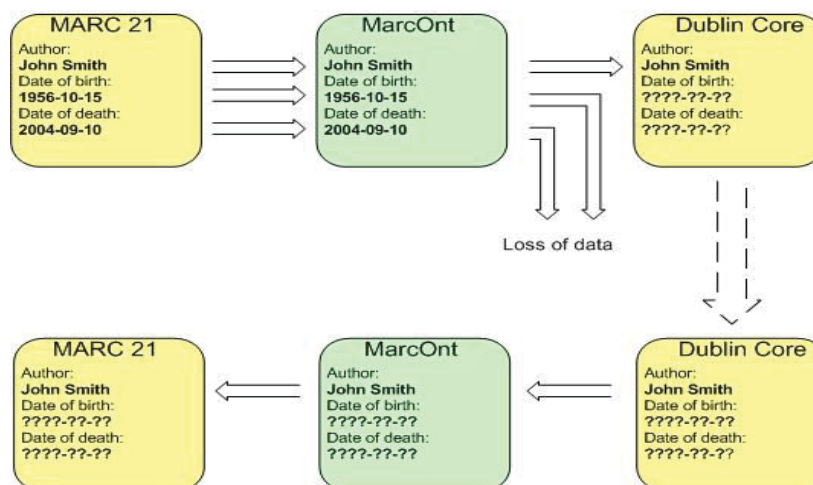
A escolha da simplicidade foi feita para a implementação inicial que poderia ser usada por algum tempo em um sistema de produção e se haveria a perda de algumas características que poderiam ser adicionadas. Esta abordagem ajuda a evitar a criação desnecessária de códigos e regras que nunca serão usados. (NOVÁCEK et al., 2007).

RDF Translator opera sobre os elementos mais básicos de grafos e triplas de RDF. Ele permite traduzir um conjunto de triplas para outro conjunto. O conjunto de triplas da entrada (ou premissas) e um conjunto de triplas de saída (ou consequências) criam uma regra simples. Basicamente, se *RDF Translator* encontra um conjunto de triplas correspondentes aos especificados nas regras das premissas, ele criará um novo conjunto de triplas de acordo com as consequências especificadas. (SYNAK, 2005; NOVÁCEK et al., 2007).

A figura 12 apresenta um mapeamento de elementos descritivos de um recurso bibliográfico nos esquemas MARC 21 e DC sem a utilização do conversor *RDF Translator*.

³⁴ *Literals* representam os tipos de dados que o valor de uma propriedade pode assumir. Os tipos mais usuais de *literals* são os do tipo *string*.

FIGURA 12: Perda de dados sem o uso de RDF.



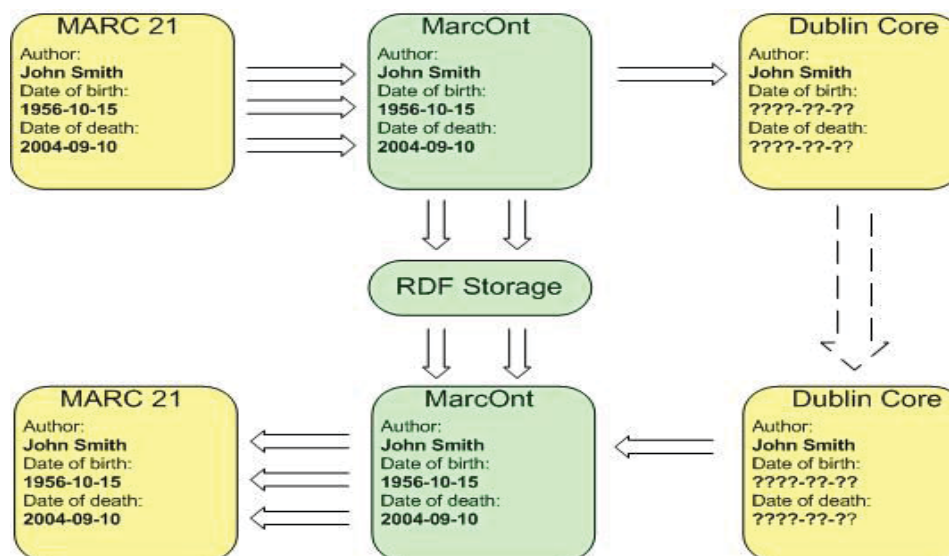
Fonte: Dabrowski e Kruk (2007, p. 11)

Pode-se observar de acordo com a figura 12, que no processo de interoperabilidade entre esquemas heterogêneos houve a perda de elementos descritivos, por exemplo, em MARC 21, os metadados descritivos tais como autor, data de nascimento e data de falecimento inseridos na entrada (*input*), foram inconsistentes na saída (*output*), tais como as datas (nascimento e falecimento). Tal fato se explica, pois não é possível converter dados de MARC 21 para MarcOnt, sem a utilização de algum mecanismo conversor.

O que na modelagem de banco de dados ou dos catálogos implicará em interfaces de apresentação dos recursos informacionais comprometidas e ainda os sistemas poderão conter possíveis inconsistências de dados no processo de busca e de recuperação de informações significativas e relevantes para os usuários.

Ao se adotar a arquitetura de metadados RDF, juntamente com o conversor *RDF Translator* os sistemas e os ambientes informacionais conseguem guardar todas as descrições bibliográficas dos esquemas de metadados heterogêneos, armazenando-as numa única interface, possibilitando dessa forma, a consistência e a integridade dos dados informacionais, oferecendo aos usuários mecanismos de busca e de recuperação de recursos com uma maior riqueza, e ainda, assegurando a interoperabilidade semântica, conforme pode ser visto na figura 13.

FIGURA 13: Adoção de RDF para a interoperabilidade semântica.



Fonte: Dabrowski e Kruk (2007, p. 12)

Ressalta-se nesse momento que na Catalogação Descritiva, todos os recursos informacionais devem estar descritos e representados na sua integridade e de maneira completa e complexa, a fim de multidimensionar as formas de acesso ao recurso, tornando-o único e individualizado, garantindo sua unicidade em um conjunto de recursos disponíveis em ambientes informacionais.

Uma vez um recurso informacional bem descrito, de acordo com normas, regras e padrões internacionais, assegurarão que os ambientes informacionais possam interoperar seus recursos, sublinhando que é o padrão de metadados que garante a interoperabilidade semântica entre os ambientes e sistemas informacionais, o que pode ser retomado na concepção primária da iniciativa MarcOnt e sua proposta de interoperabilidade semântica entre padrões de metadados heterogêneos.

Dessa forma, busca-se explicitar as principais terminologias da ontologia MarcOnt comparando-as aos códigos e esquemas de descrição da Biblioteconomia e da Ciência da Informação, em especial o AACR2 e RDA, e o formato de intercâmbio bibliográfico MARC 21.

3.4 Análise comparativa conceitual entre as principais terminologias (classes) do MarcOnt à luz dos esquemas de descrição da Biblioteconomia e Ciência da Informação (AACR2 e RDA)

A construção dessa etapa na tese, se dá principalmente para verificar as linguagens e as terminologias empregadas pela iniciativa MarcOnt, no intuito de dissipar qualquer dúvida conceitual sobre o que são os recursos (classes) num comparativo com os instrumentos de descrição na esfera da Catalogação Descritiva, com destaque para os códigos *Anglo American Cataloging Rules, 2ND edition* (AACR2) e o novo código *Resource Description and Access* (RDA).

A análise conceitual partiu da necessidade de verificar como a ontologia MarcOnt denomina cada recurso, para assim, observar as similaridades e as diferenças conceituais contempladas pelos esquemas de descrição da Biblioteconomia e Ciência da Informação, bem como apontar tais esquemas para a consolidação efetiva da iniciativa MarcOnt.

Vale a pena elucidar que as principais classes encontradas em MarcOnt, pelo menos as que se encontram conceituadas de acordo com Kruk, Synak e Zimmermann (2005c) são: Artefato, Livro, Material Cartográfico, Arquivo de computador, Recurso Contínuo, Material Misto, Música, Gravação de Som e Material Visual. Assim, a análise comparativa obedecerá tais recursos e será descrita com os termos na mesma grafia. Os quadros, em seguida, mostram a ontologia para a descrição bibliográfica, a partir da iniciativa MarcOnt e também dos códigos de catalogação AACR2 e RDA e, ainda, o padrão de metadados MARC 21.

Quadro 6: Conceitos dos recursos informacionais em MarcOnt.

MarcOnt	
Artefato	Usado para descrever todos os tipos de recursos que podem ter algum tipo de objeto sólido, como amostras médicas, maquetes, modelos, esculturas etc.
Livro	Representa todos os tipos de linguas de materiais, impressos, manuscritos etc.
Material cartográfico	Classe que pode ser usada para descrever um mapa ou um globo. Subclasses adequadas de Material Cartográfico são especificadas para acomodar descrições de diferentes recursos cartográficos.
Arquivo de computador	Significa qualquer tipo de recurso que está associado com outros computadores, incluindo software ou dados, mas também serviços e sistemas online.
Recurso contínuo	Representa todos os tipos de recursos publicados periodicamente, como revistas ou jornais, mas também revistas científicas (que não podem ser publicadas regularmente).
Material misto	Contém uma mistura de diferentes formas de materiais. Pode ser uma coleção de algum tipo (todos os tipos de materiais relativos a uma determinada pessoa ou evento) ou, por exemplo, um pacote contendo uma gravação de vídeo de uma ópera e um livreto em formato textual.
Música	É uma classe apropriada para todos os tipos de recursos musicais na forma de gravações ou notas.
Gravação de som	É qualquer coisa gravada que não é música. Contém diferentes discursos, sons e natureza.
Material visual	Contém imagens, gráficos, fotografias, projeções de filmes, mas também imagens em movimento, animações e qualquer outro tipo de material que pode ser visto.

Fonte: Elaborado pelo autor.

Quadro 7: Conceitos dos recursos informacionais em AACR2.

AACR2	
Artefato	Qualquer objeto feito ou modificado por uma ou mais pessoas
Livro	Não contemplado
Material Cartográfico	Qualquer material que representa a Terra ou algum corpo celeste, no todo ou em parte, em qualquer escala. Inclui mapas e plantas, em duas e três dimensões (incluindo mapas de lugares imaginários); cartas aeronáuticas de navegação e celestes; atlas, globos, diagramas de bloco, seções, fotografias aéreas com finalidade cartográfica, vistas panorâmicas etc.
Arquivo de computador	Material codificado (dados e/ou programa(s)) para ser manipulado por um dispositivo computadorizado. Este material pode exigir a utilização de periféricos ligados diretamente a um dispositivo computadorizado (por exemplo, um drive de CD-ROM) ou a conexão a uma rede de computadores (por exemplo, a Internet).
Recurso Contínuo	Um recurso bibliográfico que é publicado através do tempo não tendo prazo predeterminado para sua conclusão. Os recursos contínuos incluem as publicações seriadas e recursos integrados em andamento.
Material Misto	Não contemplado
Música	Obra musical que representa uma alteração de outra obra (por exemplo, transcrição livre), ou parafraseia partes de várias obras ou o estilo usual de outro compositor, ou ainda, que se baseia simplesmente em outra música (por exemplo, variações sobre um tema). Obra musical, ou parte dela, reescrita para um meio de execução diferente daquele a que visava a obra original; igualmente uma versão simplificada de obra destinada ao mesmo meio de execução.
Gravação de som	Registro de vibrações sonoras por meios mecânicos ou elétricos, de maneira a permitir a reprodução do som.
Material Visual	Não contemplado

Fonte: Elaborado pelo autor.

Quadro 8: Conceitos dos recursos informacionais em RDA.

RDA	
Artefato	Não contemplado
Livro	O resultado de uma folha impressa dobrada para formar uma reunião de folhas.
Material cartográfico	Uma representação, normalmente para escala e sobre um meio bidimensional, de uma seleção de material ou características sobre um resumo, ou em relação a, a superfície da Terra, um outro corpo celeste ou um lugar imaginário.
Arquivo de computador	Um recurso (dados e/ou programa(s)) codificado para manipulação de um dispositivo computadorizado. O recurso pode exigir a utilização de periféricos ligados diretamente a um dispositivo computadorizado (por exemplo, um drive de CD-ROM) ou a conexão a uma rede de computadores (por exemplo, a Internet).
Recurso contínuo	Não contemplado
Material misto	Não contemplado
Música	Não contemplado
Gravação de som	Uma gravação em que as vibrações sonoras foram registradas por meios mecânicos ou elétricos para que o som possa ser reproduzido.
Material visual	Não contemplado

Fonte: Elaborado pelo autor.

Quadro 9: Principais conceitos dos recursos informacionais encontrados em MARC 21.

MARC 21	
Artefato	Não contemplado.
Livro	Usado para material textual, por natureza monográfica.
Material Cartográfico	Usado para todos os tipos de material cartográfico, incluindo mapas planos e globos.
Arquivo de computador	Usado para informações codificadas que só podem ser processadas por computador ou máquinas afins, incluindo tanto os dados armazenados em formato legível por máquina, quanto os programas para processá-los.
Recurso Contínuo	Não contemplado.
Material Misto	Principalmente coleções de arquivos e manuscritos, de diversos tipos de materiais. Esses materiais podem ser de natureza monográfica ou seriada.
Música	Usado para música manuscrita e impressa, registros sonoros musicais e não musicais.
Gravação de som	Não contemplado.
Material Visual	Usado para materiais projetáveis, gráficos bidimensionais, artefatos tridimensionais ou objetos naturais e <i>kits</i> . Usado também para arquivamento de materiais visuais quando o formato ou o meio estão sendo enfatizados.

Fonte: Elaborado pelo autor.

Para melhor compreensão da explicitação da ontologia para a descrição bibliográfica faz-se a compilação dos esquemas de descrição bibliográfica e da ontologia MarcOnt, conforme pode ser observada no Quadro 10.

Quadro 10: Sistematização conceitual dos principais recursos de MarcOnt em relação ao AACR2, RDA e MARC 21.

	MARCONT	AACR2	RDA	MARC21
Artefato	Usado para descrever todos os tipos de recursos que podem ter algum tipo de objeto sólido, como amostras médicas, maquetes, modelos, esculturas etc.	Qualquer objeto feito ou modificado por uma ou mais pessoas	Não contemplado	Não contemplado
Livro	Representa todos os tipos de línguas de materiais, impressos, manuscritos etc.	Não contemplado	O resultado de uma folha impressa dobrada para formar uma reunião de folhas.	Usado para material textual, por natureza monográfica

	MARCONT	AACR2	RDA	MARC21
Material Misto	Contém uma mistura de diferentes formas de materiais. Pode ser uma coleção de algum tipo (todos os tipos de materiais relativos a uma determinada pessoa ou evento) ou, por exemplo, um pacote contendo uma gravação de vídeo de uma ópera e um livreto em formato textual.	Não contemplado	Não contemplado	Principalmente coleções de arquivos e manuscritos, de diversos tipos de materiais. Esses materiais podem ser de natureza monográfica ou seriada
Música	É uma classe apropriada para todos os tipos de recursos musicais na forma de gravações ou notas.	Obra musical que representa uma alteração de outra obra (por exemplo, transcrição livre), ou parafraseia partes de várias obras ou o estilo usual de outro compositor, ou ainda, que se baseia simplesmente em outra música (por exemplo, variações sobre um tema). Obra musical, ou parte dela, reescrita para um meio de execução diferente daquele a que visava a obra original; igualmente uma versão simplificada de obra destinada ao mesmo meio de execução.	Não contemplado	Usado para música manuscrita e impressa, registros sonoros musicais e não musicais.
Gravação de Som	É qualquer coisa gravada que não é música. Contém diferentes discursos, sons e natureza.	Registro de vibrações sonoras por meios mecânicos ou elétricos, de maneira a permitir a reprodução do som.	Uma gravação em que as vibrações sonoras foram registradas por meios mecânicos ou elétricos para que o som possa ser reproduzido.	Não contemplado

	MARCONT	AACR2	RDA	MARC21
Material visual	Contém imagens, gráficos, fotografias, projeções de filmes, mas também imagens em movimento, animações e qualquer outro tipo de material que pode ser visto.	Não contemplado	Não contemplado	Usado para materiais projetáveis, gráficos bidimensionais, artefatos tridimensionais ou objetos naturais e kits. Usado também para arquivamento de materiais visuais quando o formato ou o meio estão sendo enfatizados.

Fonte: Elaborado pelo autor.

Na análise do **MarcOnt** percebe-se que todas as classes dos recursos informacionais são definidas e conceituadas. O mesmo não ocorre em alguns tipos de recursos encontrados nos instrumentos de descrição bibliográfica. A classe livro, por exemplo, não é definida de acordo com o esquema AACR2, somente em RDA e **MarcOnt**. Outro ponto que se elucida é a diferença entre termos nos próprios esquemas de descrição bibliográfica (AACR2 e RDA), ou seja, conceitos iguais foram encontrados para termos descritos de forma diferentes.

Vale à pena destacar que as bases norteadoras das definições dos esquemas do domínio bibliográfico (AACR2 e RDA) foram extraídas dos Glossários de ambos, Apêndice D no caso do AACR2 e do formato MARC 21 para dados bibliográficos (LIBRARY OF CONGRESS, 2011).

3.5. Considerações e apontamentos sobre o **MarcOnt**

A ontologia **MarcOnt** está sendo apontada pela literatura científica como uma recomendação ou uma referência com a finalidade de fornecer uma solução semântica para a interoperabilidade no domínio bibliográfico, mais exatamente entre os esquemas e/ou padrões de metadados.

Acredita-se que a iniciativa **MarcOnt** vai ao encontro no que diz respeito à descrição bibliográfica semântica com aplicações efetivas relacionadas aos recursos bibliográficos no âmbito biblioteconômico.

De acordo com as definições das classes e das propriedades dos recursos informacionais analisadas na ontologia **MarcOnt**, e tendo como base o referencial que dá a sustentação teórica para essa investigação, pontua-se que:

Se o **MarcOnt** é uma ontologia, criada ou pensada no domínio bibliográfico, ela pode ser considerada como um esquema com orientações de descrição, que contempla os elementos de metadados (categorias individuais ou classes) que sustenta a descrição de um recurso informacional e os seus relacionamentos, apoiada no *framework* RDF, ou seja, uma base

ontológica que permite a integração de múltiplos esquemas de metadados para a descrição de recursos na *Web*, a fim de interoperar sistemas e ambientes informacionais digitais.

Pensando nessa afirmação, acredita-se que o objetivo específico número um dessa tese foi cumprido, uma vez que foi identificado, caracterizado e analisado o **MarcOnt**, objeto de estudo dessa tese, na sua essência. Não como fora denominado no projeto inicial de doutorado, enquanto um padrão de representação da informação, mas sim, respaldada na sistematização das características fundamentais extraídas da literatura, uma ontologia que descreve um recurso informacional (forma e conteúdo), seus atributos e seus relacionamentos em ambientes informacionais digitais.

A contextualização, a identificação e a caracterização do **MarcOnt** no domínio bibliográfico é de grande relevância, sobretudo por sua contribuição teórica e metodológica para o campo da Ciência da Informação, no tocante às formas de representação da informação, ao tratamento descritivo da informação e aos processos de recuperação da informação em ambientes informacionais digitais.

Nesse sentido, e sustentado pelo arcabouço teórico dessa tese, pode-se afirmar que o escopo entre a ontologia **MarcOnt** e o formato de metadados MARC 21 compreende uma vertente de estudos pautadas nos princípios e nos propósitos adotada para avaliação sistematizada.

Vale dizer que a própria criação, o desenvolvimento e a construção do **MarcOnt** tem em seu princípio a **interoperabilidade** de registros bibliográficos, oriundos de esquemas de metadados heterogêneos, construídos em diversas estruturas e pensado no domínio geral para recursos na *Web*. Em contrapartida, o padrão ou formato de metadados MARC 21 tem por objetivos **intercambiar** dados bibliográficos e catalográficos no domínio bibliográfico.

Dessa forma, torna-se inviável as interfaces de um diálogo entre **MarcOnt** e MARC 21 pelo seus próprios fundamentos e princípios; o primeiro é uma **ontologia** criada com os objetivos de interoperabilidade, enquanto o segundo, é um **formato** ou **padrão de metadados** com objetivos de intercâmbio bibliográfico.

Sendo assim, responde-se ao objetivo número três dessa tese, uma vez que não é possível relacionar e analisar as similaridades e as diferenças entre objetos que têm em sua essência, princípios e fundamentos díspares. Aponta-se a dificuldade de compreensão conceitual na literatura científica entre MarcOnt e MARC 21, sobretudo no âmbito nacional das áreas da Biblioteconomia e Ciência da Informação. Portanto, defende-se que a tese conseguiu, num primeiro momento, dissipar qualquer confusão conceitual entre o formato MARC 21 e a ontologia MarcOnt.

Um dos caminhos apontados pela literatura científica é a adoção de arquiteturas de metadados, mais especificamente do *framework* RDF, como o pilar para a iniciativa MarcOnt, além de possibilitar a interoperabilidade semântica entre ambientes informacionais digitais.

De acordo com o idealizador do MarcOnt Sebastian Kruk, a arquitetura de metadados RDF, constitui o centro ou “coração” da ontologia que aspira ser uma referência mundial, na possibilidade de funcionar como uma ferramenta para o armazenamento de descrições bibliográficas semânticas e a comunicação entre recursos informacionais de ambientes heterogêneos, com diferentes arquiteturas e plataformas.

Os ambientes informacionais digitais, especialmente o domínio particular das bibliotecas digitais, ao adotarem a arquitetura de metadados RDF, juntamente com ontologias e padrões de metadados podem encontrar um caminho para o estabelecimento da interoperabilidade semântica entre sistemas heterogêneos e distintos.

Assim, o capítulo 4 apresenta a arquitetura de metadados RDF apontada como uma promessa que garantirá a interoperabilidade em ambientes informacionais digitais e potencializará os relacionamentos das entidades bibliográficas.



CAPÍTULO 4



RESOURCE DESCRIPTION FRAMEWORK

(RDF): PROMESSA PARA INTEROPERABILIDADE

EM AMBIENTES INFORMACIONAIS DIGITAIS

4 RESOURCE DESCRIPTION FRAMEWORK (RDF): promessa para interoperabilidade em ambientes informacionais digitais

Os computadores são incrivelmente rápidos, precisos e burros; os homens são incrivelmente lentos, imprecisos e brilhantes; juntos, seu poder ultrapassam os limites da imaginação. *Albert Einstein*

Pensando na heterogeneidade de recursos informacionais distribuídos na rede e estruturados nos mais diversificados formatos de armazenamento, a palavra de ordem é marcada pela interoperabilidade, ou seja, a capacidade de ambientes e sistemas informacionais comunicarem-se entre si, potencializando os relacionamentos e garantindo uma recuperação de informações para sujeitos humanos e não humanos.

No cenário da interoperabilidade alguns modelos de dados, arquiteturas de metadados e modelos conceituais, sobretudo, no domínio bibliográfico são requisitos para garantir a modelagem de ambientes digitais altamente estruturados. Dentre eles pode-se destacar o RDF recomendado pelo W3C na construção da *Web Semântica* e sua aplicação em comunidades específicas, tais como as bibliotecas digitais e as propostas que encaminham para o delineamento de instrumentos que permitam a descrição e a representação de recursos na *Web*, como o *Resource Description and Access* (RDA). No entanto, antes de se abordar o RDA e a arquitetura de metadados RDF, será apresentado o conceito de interoperabilidade.

4.1 A visão de interoperabilidade no domínio bibliográfico

A Internet está promovendo o acesso a um número amplo, diversificado e crescente de sistemas de informação digital. Nesse contexto, as bibliotecas digitais concebidas originalmente como sistemas locais de recuperação da informação para um grupo especializado de usuários, graças à universalização da *Web* tem permitido o acesso aos recursos informacionais de forma globalizada.

A construção de bibliotecas digitais implica a integração de sistemas complexos que incluem coleções de documentos com estruturas diferentes, distintas mídias e conteúdos heterogêneos. Arelado a isso, tem-se que considerar uma variedade de componentes de *hardware* e *software*, que almejam a interoperabilidade através de diferentes estruturas de dados, algoritmos para o processamento, e ainda múltiplos sujeitos, comunidades e instituições que possuem objetivos, políticas e culturas diferenciadas.

Segundo Méndez Rodríguez (2002) a noção fundamental e o conceito chave para o funcionamento pleno da recuperação da informação em ambientes informacionais digitais estão baseados no uso efetivo dos metadados, que estabelecerá a interoperabilidade. Essa interoperabilidade requer conhecer todo o sistema de informação sob vários pontos de vista: a sintaxe (a combinação das palavras ou gramáticas para comunicar semântica e estrutura); a estrutura (que pode ser lida por humanos e pode validar-se através de máquinas); o vocabulário (que seria a correspondência entre os símbolos e os conceitos); e a semântica (o significado dos elementos ou do vocabulário).

Muitas discussões das investigações em temas que envolvem a recuperação da informação, sobretudo, em bibliotecas digitais distribuídas em rede recaem sobre a infraestrutura, em estabelecer ferramentas comuns, tecnologias e padrões compartilhados que possam facilitar a base para diversos aspectos da interoperabilidade.

Um objetivo comum desses esforços é permitir que diferentes comunidades, com diferentes tipos de informação e usando diferentes tecnologias, consigam um nível geral de compartilhamento de informação e, por meio de processos de agregação apoiados por tecnologia da informação, criem novos e mais poderosos tipos de serviços de informação. (SAYÃO; MARCONDES, 2008).

De acordo com Méndez Rodríguez (2002, p. 312, tradução nossa),

A interoperabilidade supõe complementar uma série de estratégias humanas ligadas à descrição consistente e flexível da informação mediante metadados, com um conjunto de ferramentas, *softwares* e protocolos que permitam os distintos sistemas de informação (bibliotecas digitais) estabelecer mecanismos de conversão automática entre eles, e ainda, reconciliar suas diferenças e alcançar, ao menos, um acesso universal integrado aos recursos informacionais.

A interoperabilidade, entretanto, no campo científico das áreas de Ciência da Informação e da Biblioteconomia não é algo novo, principalmente com o fenômeno da denominada explosão da informação, em meados do século XX, onde as unidades de informação, em especial as bibliotecas, sempre trabalharam numa filosofia de cooperação, colaboração e compartilhamento de recursos informacionais, criando uma rede estruturada e ordenada de informações para o acesso e a recuperação de documentos. Tais características da interoperabilidade podem ser visualizadas intrinsecamente na construção dos primeiros catálogos bibliográficos.

Na área de tecnologias da informação a interoperabilidade é vista como a capacidade de comunicação entre programas de computadores; entretanto, no campo das bibliotecas digitais o conceito de interoperabilidade é mais abstruso e estratificado e com muitas visões interdisciplinares subjacentes. (SAYÃO; MARCONDES, 2008).

A interoperabilidade pode ser considerada como o processo contínuo de assegurar que sistemas, procedimentos e a cultura de uma organização sejam gerenciados de tal forma que possibilitem a maximização das oportunidades para intercâmbio e reuso de informação. (GUY, 2005; RILEY, 2010).

Vale destacar que a interoperabilidade pode ser realizada em diferentes níveis, e possui muitas facetas, conforme pode ser observado em Guy (2005), Sayão e Marcondes (2008):

- **Interoperabilidade técnica** – as considerações sobre os aspectos técnicos incluem assegurar envolvimento de um conjunto de organizações no contínuo desenvolvimento de padrões de comunicação, transporte, armazenamento e representação de informações, tais como são o Z39.50 e o *Search Retrieval Web Service* (SRW). Inclui também os esforços cooperativos para assegurar que padrões individuais evoluam em benefício da comunidade envolvida e para facilitar, onde for possível, convergência desses padrões, de forma que seja possível que os sistemas possam ter como base mais de um conjunto de padrões.

- **Interoperabilidade semântica** – está relacionada com o significado ou semântica das informações originadas de diferentes recursos e é solucionada pela adoção de ferramentas comuns ou/e mapeáveis de representação da informação, como esquemas de metadados, classificações, tesouros e mais recentemente, ontologias; um exemplo de questão endereçada por essa faceta da interoperabilidade pode ser o seguinte: o que significa “autor” para um recurso informacional? Será a mesma coisa que “criador” para um outro recurso? No nível semântico Marino (2001) acrescenta que existem dois subníveis: Epistemológico (relacionado ao significado dos elementos descritores do formato e das relações nele existentes) e Ontológico (relacionado ao uso de ontologias, vocabulários controlados e padrões de metadados para o estabelecimento dos significados dos dados representados). Como exemplo desse nível de interoperabilidade pode-se destacar os metadados que têm o mesmo significado entre elementos de descrição do formato MARC 21.
- **Interoperabilidade política/humana** – independente das questões relacionadas à maneira pela qual a informação é descrita e disseminada, a decisão de tornar os recursos informacionais mais amplamente disponíveis e interoperáveis tem implicações para a organização, para as equipes envolvidas e para os usuários em termos comportamentais, de recursos e de treinamento. A ênfase dada por parte de alguns setores governamentais aos problemas de democratização do acesso, da exclusão digital e da federação de fontes de informação voltadas para a educação a distância, tem impacto nas políticas públicas para a área, e estão enquadrados nesse item.
- **Interoperabilidade intercomunitária** – enfoca a necessidade, cada vez mais urgente, impulsionada pela crescente interdisciplinaridade, principalmente nas áreas de pesquisa, de acesso a informações provenientes de um espectro amplo de fontes distribuídas por organizações, áreas de conhecimento e comunidades de natureza distintas. Geralmente exige o estabelecimento de fóruns para discussão e consenso em torno de práticas e procedimentos comuns.
- **Interoperabilidade legal** – considera as exigências e as implicações legais de tornar livremente disponíveis itens de informação.

- **Interoperabilidade internacional** – quando se atua em escala internacional é necessário contornar a diversidade de padrões e normas, os problemas de comunicação, as barreiras linguísticas, as diferenças no estilo de comunicação e na falta de uma fundamentação comum.

Vale destacar que essa tese tem como interesse os requisitos de interoperabilidade nas esferas semântica, estrutural e sintática, uma vez que defende-se que a estruturação e a modelagem de catálogos bibliográficos (banco de dados) ocorrem nessas camadas, para que a interoperabilidade entre ambientes informacionais digitais heterogêneos ocorra de maneira efetiva e se estabeleçam os relacionamentos bibliográficos semânticos numa mesma interface.

Méndez Rodríguez (2002) elucida que a interoperabilidade sintática (sintaxe) diz respeito à forma como se codificam os metadados digitalmente. Um exemplo dessa natureza é o RDF, que codifica os tipos de propriedades e os valores desses elementos. A estrutura descreve os mecanismos para agrupar os elementos de metadados, assim como as regras sobre como devem ser codificados os valores para cada elemento.

Metadados podem ser considerados instrumentos cada vez mais necessários no ambiente da *Web* atualmente, permitindo em larga escala a gestão de recursos informacionais distribuídos. Nos últimos anos tem-se assistido a um crescimento na interação de comunidades de metadados, que são cada vez mais impulsionadas por uma necessidade de colaboração e intercâmbio entre domínios heterogêneos.

No entanto, os padrões de metadados isoladamente não têm sido capazes de atender as necessidades de interoperabilidade entre as comunidades, instituições e organizações independentes responsáveis pelo estabelecimento da normalização e padronização de metadados.

Como podemos definir conceitualmente interoperabilidade no contexto dos metadados? O IEEE (2000) diz que a interoperabilidade corresponde à habilidade de dois ou mais sistemas ou componentes trocar informações e utilizá-las neste processo. Nilsson (2010) diz que a interoperabilidade de metadados consiste na capacidade de dois ou mais sistemas, ambientes informacionais ou componentes trocarem dados descritivos sobre recursos

informativos, e interpretá-los de uma maneira consistente, não ocorrendo falhas no intercâmbio para sua compreensão por máquinas.

De acordo com Sayão e Marcondes (2008, p. 136),

Uma organização verdadeiramente interoperável é capaz de maximizar o valor e o potencial de reuso da informação que está sob o seu controle. É também capaz de intercambiar efetivamente estas informações com outras organizações igualmente interoperáveis, permitindo que novos conhecimentos possam ser gerados a partir da identificação de relacionamentos entre conjuntos de dados previamente não relacionados. Na perspectiva do usuário, as interfaces devem apresentar para o usuário uma visão unificada em termos semânticos de diferentes recursos informativos heterogêneos, ou seja: como nomeá-los, como referenciá-los, como utilizá-los em buscas, como acessá-los, como apresentá-los para o usuário.

Pode-se dizer que a interoperabilidade é a capacidade de dois ou mais sistemas, ambientes informativos digitais e padrões de metadados heterogêneos, intercambiarem seus dados, a partir da codificação das regras e esquemas de descrição utilizadas pelo catalogador, na confecção padronizada e metodológica dos metadados, e na estruturação conceitual do ambiente informativo proporcionados pelas ontologias. Nesse sentido, Alves (2005, p. 110) diz que as ontologias funcionam como “qualificadores dos metadados e não haverá estabelecimento da semântica sem a representação por metadados”.

Os metadados são introduzidos nos sistemas de gerenciamento de informações, com o intuito de apoiar determinadas operações administrativas, incluindo a busca, apresentando resumos ou configurando interfaces. Na sua essência, os metadados criam um nível de compreensão, permitindo que os sistemas gerenciem recursos sem ter que mergulhar em seu desenvolvimento físico ou digital internos. (NILSSON, 2010).

Os metadados podem consistir de todo tipo de informação sobre um item, por exemplo, seu título, descrições textuais e classificações de assunto, para o acesso e as relações contextuais entre o item a ser descrito e outras coisas. Nesse momento é oportuno resgatar os conceitos intitulados pelos Requisitos Funcionais para Registros Bibliográficos (FRBR), onde um item pode ser entendido como um único exemplar de uma manifestação (IFLA, 1998), ou ainda, segundo as regras de catalogação do AACR2 (2005), um documento ou um conjunto de documentos sob qualquer forma física, publicado, distribuído ou tratado como uma entidade autônoma, constituindo a base de uma descrição bibliográfica única.

O núcleo de proposição dos metadados é que seu uso efetivo permite aos sistemas, às aplicações e aos usuários gerenciar e acessar itens sem nenhuma necessidade de interação direta com o item em si (LYTRAS; SICILIA, 2007). Por essa razão, a administração e o intercâmbio de metadados é a atividade central em muitos sistemas que gerenciam os objetos digitais e não digitais, tais como os sistemas de gerenciamento de conteúdo, repositórios de objetos de aprendizagem e bibliotecas (digitais ou convencionais).

Nilsson (2010) acrescenta que especificações e padrões de metadados acrescentam valores adicionais ao reduzir o limiar para o desenvolvimento de sistemas que realizam intercâmbio, reusam e combinam metadados de diferentes fontes. Um padrão comum garante uma melhor documentação e um melhor acesso às ferramentas reutilizáveis. Este é o valor central para a interoperabilidade de metadados. Percebendo o potencial inerente ao uso de metadados interoperáveis informado requer uma coordenação em grande escala entre os atores relevantes em um campo de prática e aplicação. Especificações de metadados tendem a ser concebidos para uma determinada comunidade particular, com mais ou menos itens bem definidos e descritos e cenários comuns de uso.

Neste cenário, pode-se destacar no domínio bibliográfico, metadados na forma de catalogação tem sido a questão desde as primeiras bibliotecas. Como os dados de uma biblioteca geralmente estão abertos e disponíveis ao mundo, os principais problemas na interoperabilidade de metadados começam a surgir. O desenvolvimento de um novo código de catalogação *Resource Description and Access* (RDA) aparece nesse contexto, como uma forma de auxiliar a interoperabilidade de metadados das bibliotecas, destacando a situação complexa com uma multiplicidade de padrões de metadados utilizados em bibliotecas do mundo todo, como por exemplo, o MARC 21.

Tanto as bibliotecas como as ferramentas tecnológicas que permeiam seu ambiente informacional atingem o campo dos metadados orientados a *Web*, onde o *Resource Description Framework* (RDF) tem feito progresso ao longo das últimas décadas, juntamente com uma série de especificações que sustentam a *Web Semântica*, tal como a *Web Ontology Language* (OWL). (WORLD WIDE WEB CONSORTIUM, 2009).

Quando os metadados projetados de acordo com diferentes especificações e de diferentes domínios se encontram, ou seja, quando comunidades se envolvem para aumentar

sua interação, dificuldades consideráveis no gerenciamento dos metadados tendem a aparecer, pois muitas vezes os dados se apresentam incompatíveis. O resultado é que os benefícios da interoperabilidade de metadados dentro de um padrão são perdidos quando são combinados, aumentam os custos de desenvolvimento e falhas na comunicação entre os sistemas, aplicando-se em muitos casos soluções não reutilizáveis.

Godby, Smith e Childress (2003) argumentam baseados em experimentos com o cruzamento de metadados (*crosswalks*), que as traduções completas de um padrão para outro são possíveis numa comunidade de prática específica, enquanto somente traduções parciais são possíveis entre elas. Os autores exemplificam os metadados de uma biblioteca que podem ser combinados com outras especificações de bibliotecas diferentes, desde que usando um padrão comum para a descrição dos recursos bibliográficos.

Nesse momento pode-se corroborar com as autoras Santos e Corrêa (2009) ao mencionarem que nos ambientes de manipulação de dados bibliográficos, em especial, os ambientes informacionais digitais, quem define a interoperabilidade é o padrão de metadados a ser adotado para representar e descrever os recursos informacionais.

Subjacentes aos formatos e modelos abstratos é a noção sutil de semântica. Com a ascensão do RDF e a iniciativa da *Web Semântica* pelo W3C, a descrição semântica de metadados tem recebido especial atenção, no desenvolvimento e na construção de ambientes informacionais digitais interoperáveis.

4.2 Resource Description Framework (RDF): o *core* para descrições bibliográficas

RDF é também uma linguagem geral proposta para a representação da informação na *Web*. Ela foi originalmente criada nos anos de 1990, tal como um padrão baseado em XML para codificação de metadados, particularmente sobre recursos *Web*.

A arquitetura de metadados RDF possibilita o processamento e o gerenciamento automatizado de recursos informacionais, de acordo com áreas do conhecimento distintas, principalmente com relação: à descoberta de recursos informacionais, à descrição das relações

entre os recursos informacionais representadas na rede, ao auxílio aos agentes de *software*, ao intercâmbio e ao compartilhamento de informações entre outros tipos de aplicações. (LASSILA; SWICK, 1999).

O RDF é uma recomendação do *World Wide Web Consortium* (W3C) para a descrição semântica de recursos. A especificação do modelo e sintaxe do RDF fornece uma sintaxe e um modelo para a codificação em XML de vocabulários como é o caso do *Dublin Core*. O RDF permite a interoperabilidade entre aplicações, e a serialização em XML é uma das mais utilizadas devido, não só a promover essa interoperabilidade, mas também ao fato da XML ser cada vez mais utilizada para transportar informação. (ARMS, 2002).

RDF é um *framework* para metadados que possibilita a interoperabilidade. Ele provê uma infraestrutura que possibilita a codificação, o intercâmbio, o uso e o (re)uso de metadados de uma forma não ambígua, de maneira que máquinas possam “entender” a semântica dos metadados e conseqüentemente usá-los no processo chamado descoberta do recurso. (TAYLOR, 2004; ZENG; QIN, 2008).

O RDF é desenhado para facilitar aos agentes de busca perceber o suficiente sobre um *Website*, de modo a que possa descobrir recursos, catalogar o conteúdo do *site*, escalonar esse conteúdo, perceber que possui o conteúdo e sob que termos e a que preço este pode ser usado, e fazer outras coisas que um *Webspider*³⁵ ou um agente inteligente possa querer fazer. (ARMS, 2002).

De acordo com Arms (2002), o RDF está dividido em duas partes, compreendendo duas especificações diferentes:

- O *RDF Model and Syntax Specification* (RDFMSS) que é uma recomendação do W3C e que apresenta um modelo para representar metadados RDF, assim como uma sintaxe para codificar e transportar metadados, de forma que maximize a interoperabilidade de servidores e clientes *Web* desenvolvidos independentemente;

³⁵ *Webspider* são agentes de *software* que passam pela coleta, filtragem e, potencialmente, agregação de informações da Internet para usuários em vários sistemas de informação. Seu objetivo pode ser coletar informações ou entender a estrutura e validade de um *Website*. *Spiders* são a base para os mecanismos de procura modernos, como Google e AltaVista. Esses *spiders* recuperam automaticamente os dados da *Web* e os transmitem para outros aplicativos que indexe os conteúdos do *Website* para um melhor conjunto de termos de busca.

- O *RDF Schema Specification* que é uma recomendação do W3C, e define uma linguagem de especificação de esquemas. Com o *RDF Schema*, pode-se desenhar e implementar de uma forma consistente, vocabulários de metadados específicos. Estes podem, ainda, ser mais desenvolvidos no seio de outros projetos gerando, assim, uma rede de esquemas de metadados. Por exemplo, os termos de um vocabulário a ser construídos podem perfeitamente ser definidos como refinamentos de elementos do *Dublin Core*, ou de outro qualquer vocabulário anteriormente definido.

O principal objetivo do RDF é facilitar o intercâmbio de informações (que podem ser interpretadas por máquinas) entre aplicativos via *Web*. Um dos atributos importantes do RDF é sua habilidade de automatizar o processamento de recursos *Web*. Essa automação refere-se ao fato de recursos *Web* tais como URL (*Uniform Resource Locator*) e URI (*Uniform Resource Identifier*), entre outros, estarem formalmente descritos em RDF. (ARMS, 2002).

A meta em RDF é muito mais precisa do que é possível em uma busca de texto na íntegra (pesquisa de texto completo a ser o método atualmente invocado pela maioria dos motores de busca). Mais precisão é possível porque RDF identifica o “*namespace*” do qual extraiu sua semântica. A máquina ao ler as declarações (*statements*) pode então interpretar adequadamente as declarações de referência para o *namespace* por meio do RDF.

RDF prove um mecanismo para a integração de múltiplos esquemas de metadados para a descrição de recursos na *Web*. RDF é um modelo de dados que prove uma estrutura independente de comunidades, pode desenvolver vocabulário que adaptam suas necessidades específicas e então compartilham aqueles vocabulários com outras comunidades.

O RDF na concepção de Arms (2002) funciona com todos os tipos de dados eletrônicos, e possui uma variedade de aplicações, entre as quais pode-se destacar:

- A identificação consistente de direitos autorais, marcas registradas e outras informações de propriedade intelectual em recursos da *Web*;
- A catalogação avançada de recursos e seus relacionamentos dentro de um único sistema ou entre vários sistemas;
- A pesquisa mais especializada de arquivos de dados, pois os dados em si estão mais bem identificados, representados e descritos.

A arquitetura de metadados RDF é composta por três tipos de objetos:

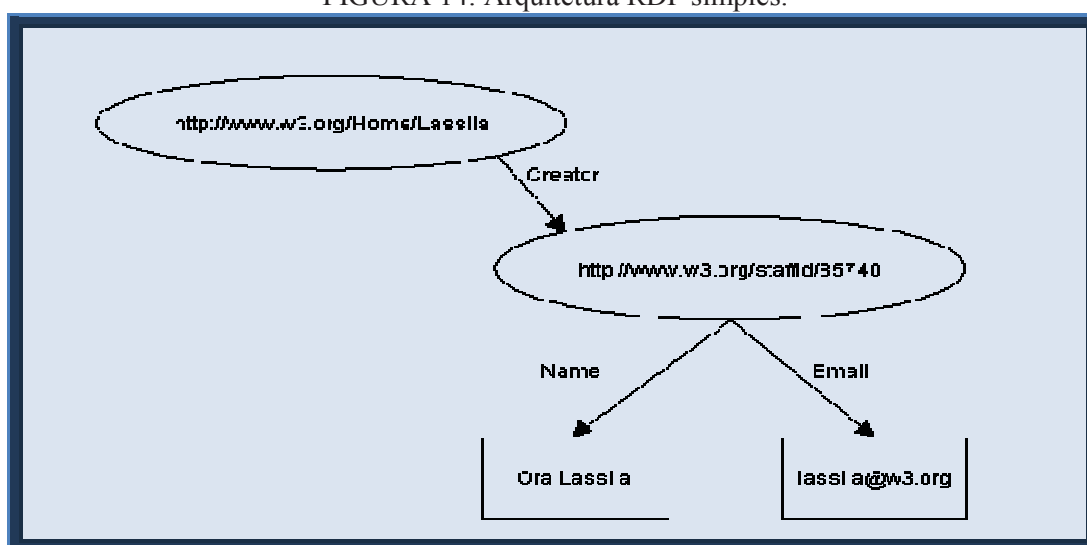
- **Recurso:** é o que será descrito por uma expressão RDF. Todo recurso é identificado por um URI (incluindo aí o URL).
- **Propriedade:** é qualquer característica utilizada para descrever um recurso. Em RDF um domínio de conhecimento é definido via um RDF *Schema*. É em um RDF *Schema*, portanto, que é definida a semântica e as características de uma propriedade. Uma aplicação que crie metadados em RDF, e outra que utilize estes metadados devem utilizar o mesmo *Schema* para um funcionamento adequado.
- **Tripla:** é formada por um recurso, uma propriedade e um valor para a propriedade daquele recurso. Uma tripla possui a seguinte forma: <sujeito, predicado, objeto>. O significado de uma tripla pode ser resumido como o recurso (sujeito) possui a propriedade (predicado) com este valor (objeto). Um valor ou objeto pode ser tanto um outro recurso quanto um tipo primitivo definido por XML.

O *parser* RDF é responsável por ler e verificar a sintaxe RDF transformar o código escrito na sintaxe RDF num conjunto de triplas e, eventualmente, num grafo RDF. Este não passa de uma representação gráfica desse conjunto de triplas: um grafo em que cada propriedade ou predicado é representado por um arco. Depois, só restam os sujeitos e os objetos que, sendo recursos, podem também ser sujeitos de uma outra propriedade, e assim sucessivamente. (ARMS, 2002).

De acordo com Zeng e Qin (2008) o maior uso de RDF não se restringe à codificação de informações sobre recursos na *Web*: também prove informações sobre, e as relações entre coisas e o mundo real: pessoas, lugares, conceitos etc. O modelo de metadados RDF é baseado no princípio de declarações lógicas (*making*) sobre recursos na forma de expressões *subject-predicate-object* (chamados tripla na terminologia RDF). Os termos RDF para as várias partes da declaração (*statements*) são:

- The *subject* (é a URL: <http://www.slis.kent.edu/~mzeng/metadatabasics/index.html>).
- The *predicate* (é a palavra título)
- The *object* (é a frase *Metadata Basics*)

FIGURA 14: Arquitetura RDF simples.



Fonte: Lassila e Swick (1999)

Conforme pode ser observado na figura 14, apresenta-se uma estrutura RDF, onde se podem traçar os seguintes comentários: o *assunto* (recurso) é definido neste diagrama pelo nó <http://www.w3.org/Home/Lassila>, cujo *predicado* (propriedade) é definido por um criador (*creator*), possuindo como objeto (*Ora Lassila*). Além disso, o recurso possui atributos e valor estruturado com identificador (85740), criado por (*Ora Lassila*) que possui o e-mail lassila@w3.org (objeto).

A arquitetura de metadados RDF pode ser representada e apresentada também na linguagem XML, conforme pode ser visto na figura 15.

FIGURA 15: Declaração RDF utilizando a sintaxe XML.

```

<rdf : RDF>
  < rdf: Description about= "http://www.w3.org/Home/Lassila" >
    <creator> Ora Lassila </ Creator >
  </ rdf: Description >
</rdf : RDF>

```

Fonte: Lassila e Swick (1999).

Conforme pode ser observado pela figura 15, uma declaração RDF utilizando a linguagem XML pode ser representada de acordo com a etiqueta *rdf : RDF* cujo conteúdo dessa etiqueta é expressa por outra etiqueta do tipo *rdf: Description*, bem como a validação ou referência de um recurso já existente, representado pela etiqueta *about*.

A arquitetura RDF em seus objetivos visa

[...] ter um modelo de dados simplificado; possuir semântica formal e provável inferência; utilizar um vocabulário extensível; utilizar sintaxe baseada em XML; dar suporte aos tipos de dados do XML Schema; permitir que qualquer pessoa possa fazer declarações acerca de qualquer recurso informacional. (FEITOSA, 2006, p. 101-102).

RDF usa a URI³⁶ assim como seus mecanismos básicos para a identificação de assuntos, predicados e objetos nas declarações (*statements*). Objetos em declaração RDF podem ser juntas *URIrefs* (ex.: o criador da URI), ou valores constantes (chamado *literal*) representado por caracteres *strings* (ex.: a data e a língua), na ordem para representar diferentes tipos de valores das propriedades. (ZENG; QIN, 2008).

RDF fornece um modelo que é independente de sintaxe. O modelo diz que um recurso é qualquer objeto que pode ser unicamente identificado por um URI. Recursos têm propriedades, e propriedades têm valores. Um valor pode ser uma sequência de caracteres ou um número, ou pode ser outro recurso. Se um valor é outro recurso, daquele recurso, por sua vez têm propriedades que têm valores, algumas das quais poderiam ser outros recursos. Não

³⁶ *Uniform Resource Identifier (Identificador Uniforme de Recursos)* é uma cadeia de caracteres compacta usada para identificar ou denominar um recurso na Internet. O principal propósito desta identificação é permitir a interação com representações do recurso através da Internet usando protocolos específicos. URIs são identificadas em grupos definindo uma sintaxe específica e protocolos associados.

há limites para o número de níveis de conexões para recursos. A coleção de propriedades de um recurso é chamada de uma descrição. (ZENG; QIN, 2008).

Descrevendo um recurso com metadados permite que ele possa ser compreendido por humanos e máquinas de forma que promova a interoperabilidade. Utilizando esquemas de metadados definidos, compartilhando transferência entre protocolos, e *crosswalks* entre esquemas, os recursos na rede podem ser buscados mais amplamente.

A coisa mais elementar que é necessário para um modelo de conceitualização compartilhada de algum domínio é a maneira de denotar entidades e referenciá-las de forma inequívoca. Nesse sentido, a arquitetura de metadados RDF em *MarcOnt* usa URIs (*Uniform Resource Identifier*) e identificadores de recursos que devem ser globalmente únicos.

Cada ontologia deve fornecer os significados para designar entidades. Existindo entidades que podem ser referenciadas, o próximo passo é descrever as relações entre elas. Como as relações são elementos fundamentais semânticos, devem também ser inequivocamente endereçável. As propriedades em RDF podem ser vistas como relações binárias. Este é o tipo mais básico de relações entre duas entidades. Os tipos de relações mais complexas podem ser modelados através da definição de um vocabulário para as propostas no topo de RDF, tal como tem sido feito em OWL. (HEATH; BIZER, 2011).

Os dois elementos para a modelagem semântica, mecanismos para identificar entidades e o estado dos relacionamentos entre eles são fornecidos por RDF. As linguagens de ontologia que são construídas em RDF utilizam estes mecanismos e definem as semânticas de certos relacionamentos, entidades e associações de relacionamentos e entidades. Dessa maneira, RDF fornece a estrutura em que as primitivas semânticas das linguagens de ontologia estão inseridas. Isso significa que se podem distinguir aqui três camadas: sintática (XML), uma camada estrutural (RDF) e uma camada semântica (linguagem de ontologia). (DABROWSKI; KRUK, 2007).

As várias linguagens de ontologia diferem no seu vocabulário, seus fundamentos lógicos e nos seus elementos epistemológicos, mas elas têm em comum o fato de descrever as estruturas das entidades e suas relações. Portanto, RDF é o maior denominador comum de todas as linguagens de ontologia. (HEATH; BIZER, 2011). RDF não é apenas uma forma de

codificar as linguagens de ontologia ou apenas um modelo de dados arbitrário, mas é um modelo de dados estruturado que corresponde exatamente à estrutura das linguagens de ontologia e à codificação de elementos de metadados.

Nesse momento é oportuno ressaltar o trabalho integrado entre RDF e outras tecnologias para a efetivação da *Web Semântica* e para que ocorra a interoperabilidade semântica entre os sistemas e ambientes informacionais digitais, dentre elas pode-se destacar *RDF Schema (RDFS)*.

Para Baca (2008) o *RDF Schema* pode ser considerado como um conjunto de semântica dentro de um *namespace* definido para uso com aplicações específicas de RDF.

O RDFS apresenta-se nesse contexto como uma linguagem de descrição de vocabulário, ou seja, descreve as propriedades de um recurso informacional, bem como estabelece as relações entre as propriedades de outros recursos, podendo, além disso, descrever as classes, as propriedades e outros recursos informacionais condizentes aos pares (atributo-valor) contemplados em XML.

O *RDF Schema*, nesse sentido, pode ser compreendido como uma ferramenta que permite descrever vocabulários, bem como as classes e as propriedades para um recurso RDF, fornecendo assim, uma estrutura mais flexível e com uma maior amplitude semântica na descrição dos recursos informacionais.

Para Heath e Bizer (2011), RDFS é uma linguagem para descrever ontologias em RDF, que são muitas vezes referidas como vocabulários. Os vocabulários consistem em definições de classes e de tipos de propriedades, por exemplo, uma classe livros, com propriedades sobre o número de páginas, o volume etc.

Existem duas classes básicas na linguagem RDFS:

- `rdf:Class`: que é a classe de recursos que são as classes RDF;
- `rdf:Property`: que é a classe de todas as propriedades RDF.

Um recurso é declarado para ser uma classe, escrevendo-a como uma instância de `rdfs:Class` usando o predicado `rdf:type`.

Em RDF, os recursos podem ser divididos em grupos chamados *classes*. Os membros de uma classe são denominados *instâncias*. As classes em si, são também recursos e podem ser descritas por meio de *propriedades*. Ao conjunto de instâncias de uma classe denomina-se *extensão da classe*. Uma classe pode ser membro de sua própria extensão, isto é, pode ser uma instância de si própria. Todas as classes em RDF Schema são agrupadas em uma classe principal denominada *rdfs:Class*. De tal maneira, se uma classe X é uma *subclasse* de Y, todas as instâncias de X serão também instâncias de Y e poderão ser declaradas com o uso da propriedade *rdfs:subClassOf*. O termo superclasse é utilizado para designar uma classe superior em relação a suas subclasses. Coleções de recursos podem ser representadas em RDF pelo uso de *contêineres*, que podem ser de valores alternativos (*rdf:Alt*), de valores não ordenados (*rdf:Bag*) e de valores sequenciais (*rdf:Seq*). (HEATH; BIZER, 2011; FEITOSA, 2006).

Para potencializar ainda mais as formas de busca e de recuperação das informações nos ambientes informacionais digitais e promover a interoperabilidade semântica entre os sistemas e ambientes, o W3C recomendou e desenvolveu o RDFa, que de acordo com Adida et al. (2008) seria uma especificação de atributos para expressar dados estruturados em qualquer linguagem de marcação.

A essência do RDFa³⁷ é generalizar os atributos (daí o nome RDFa) dos elementos dos metadados para permitir o uso em qualquer elemento num documento XHTML³⁸. Para Adida et al. (2008), os atributos podem ser assim entendidos:

³⁷ Maiores informações sobre o RDFa podem ser encontrados no *site*: Disponível em: <<http://www.w3.org/TR/2008/REC-rdfa-syntax-20081014/>>. Acesso em: 22 maio de 2010.

³⁸ O XHTML, ou eXtensible Hypertext Markup Language, é uma reformulação da linguagem de marcação HTML, baseada em XML. Combina as *tags* de marcação HTML com regras da XML. O XHTML consegue ser interpretado por qualquer dispositivo, independentemente da plataforma utilizada, pois as marcações possuem sentido semântico para as máquinas.

#about#	uma URI especificando o recurso de que se trata o metadado; na sua ausência, por omissão, o documento corrente.
#rel#, #rev# e #href#	especifica a relação ou a relação inversa com outro recurso.
#property#	especifica uma propriedade para o conteúdo de um elemento.
#content#	atributo opcional que ultrapassa o conteúdo do elemento quando se usa o atributo property.
#datatype#	atributo opcional que especifica o tipo de conteúdo.

Vale dizer que muitos esforços têm sido despendidos na criação e no desenvolvimento de ferramentas e de instrumentos tecnológicos que podem contribuir para as questões de representação e descrição de recursos informacionais em ambientes digitais e ainda, permitir a troca efetiva, ou seja, a interoperabilidade semântica, garantindo o uso e o (re)uso da informações no atendimento às solicitações dos usuários.

No domínio bibliográfico, comunidades têm investigado orientações que podem ser potencialmente utilizadas em harmonia com o RDF, tal como o *Resource Description and Access* (RDA), projetado para a representação e a descrição de recursos informacionais em meio digital e compatível com a *Web Semântica*.

4.3 RDA: proposta para um código internacional de catalogação aplicado em ambientes digitais

Resource Description and Access (RDA) é um novo padrão para descrever metadados de recursos mantidos nas coleções de bibliotecas, arquivos, museus e outras organizações de gerenciamento de informações. Desde 2007, o grupo de trabalho do padrão *Resource Description and Access* (RDA) tem realizado um esforço paralelo para construir vocabulários habilitados no contexto da *Web Semântica*.

Seu objetivo é fornecer um conjunto compreensível de diretrizes e instruções descritivas para a criação de metadados oriundos de todos os tipos de mídias e conteúdos de recursos. Suas orientações estão baseadas no código de catalogação AACR2 e no desenvolvimento dos Requisitos Funcionais para Registros Bibliográficos (FRBR) e também na Declaração Internacional de Princípios de Catalogação, formulada na primeira década do século XXI.

De acordo com Hillmann et al. (2010), o RDA padroniza essencialmente como o conteúdo dos metadados é identificado, transcrito e estruturado, embora seja independente de qualquer codificação de metadados específica. RDA também identifica um conjunto geral de elementos de metadados, e em muitos casos, fornece um vocabulário controlado para uso, assim como o conteúdo de um elemento.

Embora RDA esteja sendo desenvolvido principalmente para uso de recursos informacionais em ambientes de biblioteca, pesquisas e consultas foram realizadas em outras comunidades de gerenciamento de informações, incluindo editores e nas que operam no ambiente informacional digital, para tentar garantir o alinhamento efetivo de padrões de metadados utilizados nessas comunidades. (HILLMANN et al., 2010).

Dentre as comunidades mencionadas por Hillmann et al. (2010), destaca-se a *Dublin Core Metadata Initiative* (DCMI). O desenvolvimento da parte de orientação descritiva do RDA tem ocorrido em paralelo com o esforço de criação de representações formais do conjunto de elementos e valores de vocabulários para o uso de sujeitos humanos e não humanos (máquinas).

Este trabalho iniciou-se como resultado de uma reunião realizada na *British Library* em abril e maio de 2007, que incluiu representantes do *Joint Steering Committee for Development of RDA* (JSC) e *Dublin Core Metadata Initiative* (DCMI). Foram incluídos também os membros da comunidade DCMI com experiência em tecnologias da *Web Semântica* e Gordon Dunsire da *University of Strathclyde*, na Escócia, representando o *RDA Outreach Group*.

Desse encontro, o qual definiu o desenvolvimento imediato de toda gama de vocabulários RDA, cresceu o grupo de trabalho DCMI/RDA, co-presidido por Diane

Hillmann e Gordon Dunsire. O grupo de trabalho levantou fundos da *British Library* e *Siderean Software* para contratar consultores e recrutar voluntários para o projeto, e começou a usar o NSDL Registry³⁹ para desenvolver e registrar os vocabulários.

A carta do grupo de trabalho define os componentes do projeto do padrão RDA, assim como um vocabulário RDF para uso no desenvolvimento de um perfil de aplicação Dublin Core. Um perfil de aplicação é definido pelo DCMI⁴⁰ (2001, tradução nossa) como,

Um conjunto de elementos de metadados, políticas e diretrizes definidas para uma aplicação em particular. Os elementos podem ser de um ou mais conjuntos de elementos, permitindo assim que uma determinada aplicação atenda aos requisitos funcionais, utilizando metadados de muitos conjuntos de elementos definidos localmente. Por exemplo, uma determinada aplicação pode escolher um subconjunto de Dublin Core que atenda suas necessidades, ou podem incluir elementos de DC, outros conjuntos de elementos, e ainda vários elementos definidos localmente, todos combinados em um único esquema. Um perfil de aplicação não é completo sem a documentação que define as políticas e as melhores práticas apropriadas para a aplicação.

Como RDA foi projetado para a representação de recursos informacionais digitais, muitas iniciativas têm acontecido para que suas orientações sejam utilizadas e reaproveitadas pelas arquiteturas de metadados, como RDF, a fim de potencializar a estruturação e o desenvolvimento de ambientes informacionais digitais na garantia da interoperabilidade.

4.4 O papel do RDF nos esquemas e modelos bibliográficos

RDF é um grupo de especificações desenvolvidas pelo *World Wide Web Consortium* (W3C), como um modelo para a descrição de metadados na rede. É uma linguagem para a representação de informações sobre recursos na *Web*.

O conceito de um recurso é generalizado em RDF para significar qualquer coisa que possa ser descrita com metadados. Isso permite aos metadados a ser aplicados a qualquer coisa sua identificação, mesmo que ela não possa ser diretamente recuperada na *Web*.

³⁹ As informações sobre os vocabulários RDA podem ser encontradas no *site*. Disponível em: <<http://rdvocab.info/>>. Acesso em: 14 nov. 2011.

⁴⁰ As definições sobre o perfil de aplicação do DCMI podem ser localizadas no seu Glossário. Disponível em: <<http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml>>. Acesso em: 14 nov. 2011.

Hillmann et al. (2010) aponta que RDF é, portanto, compatível com o atual cenário do mundo real que RDA se destina a abordar, onde os metadados são em grande parte legíveis por máquinas e os recursos descritos, por sujeitos humanos.

Hillmann et al. (2010) elucida que RDF é uma escolha adequada para aplicação em vocabulários RDA por muitas razões.

- ✓ O modelo RDF é baseado nas mais simples estruturas de metadados, uma única declaração sobre uma única propriedade de um só recurso. Tais declarações podem ser agregadas de forma flexível para formar descrições de alto nível, ou registros de metadados, de um recurso específico.
- ✓ Construir vocabulário para as entidades RDA em RDF cumpre-se o objetivo de fornecer instruções RDA para o registro de metadados que podem ser aplicados independentemente da estrutura ou sintaxe para o armazenamento e a apresentação dos dados.
- ✓ RDF assume um mundo aberto onde o armazenamento e a manutenção de metadados são distribuídos, ou seja, seus conteúdos são destinados ao compartilhamento e não mais organizados em "silos" fechados. Essa abertura é essencial para que RDA rompa com os limites das práticas das bibliotecas convencionais/tradicionais.

Em relação ao uso de ontologias e das linguagens de ontologias, Hillmann et al. (2010) explica que quando se expressa os elementos de RDA em uma maneira compatível com a *Web Ontology Language* (OWL) e a linguagem RDF *Schema* (RDFS) propicia-se suportar as regras de inferências tão importantes no mundo RDF, e permite RDA beneficiar-se da extensibilidade do modelo RDF.

Da mesma forma, valores de vocabulários em *Simple Knowledge Organization System* (SKOS⁴¹) e um vocabulário RDF construído em OWL, permitem que os vocabulários sejam estendidos e usados de acordo com as intenções almejadas em RDA. O uso de SKOS permite

⁴¹ Sistema de organização simples do Conhecimento (SKOS) é uma família de linguagens formais criados para a representação de tesouros, esquemas de classificação, taxonomias, sistemas de cabeçalhos de assunto, ou qualquer outro tipo estruturado de vocabulário controlado. SKOS é construída sobre RDF e RDFS, e seu objetivo principal é permitir a fácil publicação de vocabulários controlados estruturados para a *Web Semântica*. Atualmente SKOS é desenvolvido pelo grupo de trabalho e especificações do W3C. Maiores informações em: Disponível em: <<http://www.w3.org/TR/skos-reference/skos.html#semantic-relations>>. Acesso em: 20 jan. 2012.

também ao RDA ser mais bem integrado ao desenvolvimento geral de ontologias e organização do conhecimento que são importantes para melhorar as aplicações de recuperação da informação centradas nos usuários.

RDF requer o uso de identificadores facilmente processáveis pelas máquinas para entidades estruturais e de conteúdo. Esses identificadores são independentes de considerações da linguagem humana e permitem os vocabulários construídos em RDA ser traduzidos para diferentes linguagens, sem a necessidade de identificadores distintos. Esta é uma vantagem significativa para incentivar o uso de RDA além da comunidade anglófona.

Metadados expressos em RDF podem ser mais facilmente processados e tratados para garantir a semântica e a veracidade dos seus conteúdos, sem se deter ou prender-se na validação do seu formato ou sintaxe, e é significamente diferente da "todo mundo deve usar o mesmo esquema ou modelo de XML". (HILLMANN et al., 2010, p. 2, tradução nossa).

Baseando-se na citação anterior, pode-se dizer que isto é importante em ambientes informacionais digitais, onde a geração de metadados heterogêneos por sujeitos humanos não preparados e treinados e também por não humanos (máquinas) é uma constante.

De acordo com o editor do RDA, Tom Delsey (2010), a implementação do RDA pode ser aplicada às estruturas de vários bancos de dados, estruturando seus metadados de acordo com o vocabulário (regras de conteúdo) determinado em RDA.

Para tanto, as seções seguintes serão baseadas no pensamento de Hillmann et al. (2010), por suas relevantes contribuições no campo da Catalogação Descritiva, sobretudo, nas áreas relacionadas ao desenvolvimento, ao uso e à aplicação dos metadados, em ambientes digitais, além de fazer parte dos membros da *Dublin Core Metadata Initiative* (DCMI).

Segundo Hillmann et al. (2010) três cenários são descritos nesse contexto:

1. Um banco de dados relacional ou orientado a objeto;
2. Um banco de dados bibliográficos ligados e registros de autoridades;
3. E um banco de dados de arquivo simples.

RDA é otimizado para o primeiro e segundo cenários, onde a estrutura, ou seja, a modelagem do banco de dados é implementada segundo os modelos conceituais do FRBR e dos Requisitos Funcionais para Dados de Autoridade (FRAD). Isso resulta em metadados para um único recurso a ser distribuídos através de registros que contém dados para a descrição e o acesso, com a duplicação reduzida de dados e um melhor potencial para o (re) uso. O Grupo de Trabalho assumiu que esta abordagem será desenvolvida, por exemplo, pelo tratamento de valores de vocabulários RDA como “arquivos de autoridade”. Esta decomposição *top-down* de registros de catálogos monolíticos é compatível com uma agregação *botton-up* de declarações RDF. (HILLMANN et al., 2010).

Para Hillmann et al. (2010), RDA precisa ser compatível com as práticas atuais, a fim de incitar a adesão, e a mais prevalente implantação de banco de dados nas bibliotecas de hoje são aquelas categorizadas por registros bibliográficos e de autoridade que são melhores exemplificados pelo padrão ou formato de metadados MARC 21. RDA procura ser compatível com os três cenários, mas isso dificulta a extensão do cenário de banco de dados relacional em RDF puro para representações RDA, o que implica numa série de desafios pelo Grupo de Trabalho RDA.

Um aspecto significativo do RDA é o uso do modelo entidade-relacionamento dados em FRBR. O uso do modelo entidade-relacionamento (ER) é um componente-chave da *Web Semântica*, mas é inteiramente novo nas regras de catalogação em bibliotecas. As revisões do RDA realizadas pelo JSC, ao longo de seu desenvolvimento, fizeram o comitê repensar as bases conceituais tradicionais estabelecidas pelas regras de catalogação. É importante destacar que o desenvolvimento dos elementos de RDA reflete o comprometimento do JSC junto aos princípios dos FRBR, embora a interação de RDA com os FRBR não tenha sido desenvolvida de tal forma que traduziria diretamente a correspondência dos relacionamentos em RDF. (HILLMANN et al., 2010). Nesse sentido, o grupo DCMI/RDA tem trabalhado no desenvolvimento e no registro de vocabulários para sanar as principais lacunas que possam existir na correspondência dos elementos RDA nas representações em RDF.

A integração entre RDA e FRBR significava que RDA em RDF seria necessário para interagir com um RDF compatível ao tratamento de FRBR, preferencialmente com um acordo oficial entre os seus criadores. No entanto, a IFLA foi incapaz de fornecer as entidades FRBR com os requisitos e identificadores necessários estruturados em RDF em tempo hábil. Além

disso, o trabalho da IFLA sobre FRAD só recentemente chegou ao seu fim. Portanto, para realizar a criação de propriedades RDA em RDF foi necessário criar uma versão transacional de FRBR, preferencialmente no mesmo registro de RDA. Esta versão específica RDA de FRBR inclui “Família”, dos FRAD, e “Agente” do modelo orientado a objeto de FRBR, e inclui apenas o nível da entidade, com as entidades identificadas como classes e subclasses. Uma versão RDF de FRBR está oficialmente disponível a partir da IFLA, os relacionamentos entre as mesmas classes definidas na versão RDA de FRBR e uma versão IFLA dos FRBR será feita para indicar que estas são realmente as mesmas entidades. (HILLMANN et al., 2010).

Fazer a conexão entre as entidades FRBR e as propriedades RDA não foi simples (DELSEY, 2010). O primeiro passo foi definir elementos RDA como propriedades RDF, tentando atribuir cada propriedade para uma e somente uma entidade específica de FRBR. Havia preocupações significativas com essa abordagem, porque os elementos de RDA eram limitados e não correspondiam como os FRBR poderiam ser implementados na prática.

Comunidades especializadas com uma visão diferente e de acordo com as necessidades de seus usuários não teriam outra alternativa do que criar novos elementos de dados (com diferentes relacionamentos para entidades FRBR) para expressar seus metadados.

De acordo com Hillmann et al. (2010), os estudos realizados por bibliotecários de música e catalogadores de audiovisuais mostram que a definição de entidades do Grupo 1 dos FRBR é altamente diferenciada quando se lida com diferentes formas de expressão criativa. Além disso, há elementos de RDA que estão vinculados explicitamente às orientações de RDA para uma entidade FRBR, geralmente em virtude da inclusão da entidade em nome do elemento. Exemplos destes últimos são elementos do tipo “Identificador da Manifestação” e “Linguagem da Expressão”.

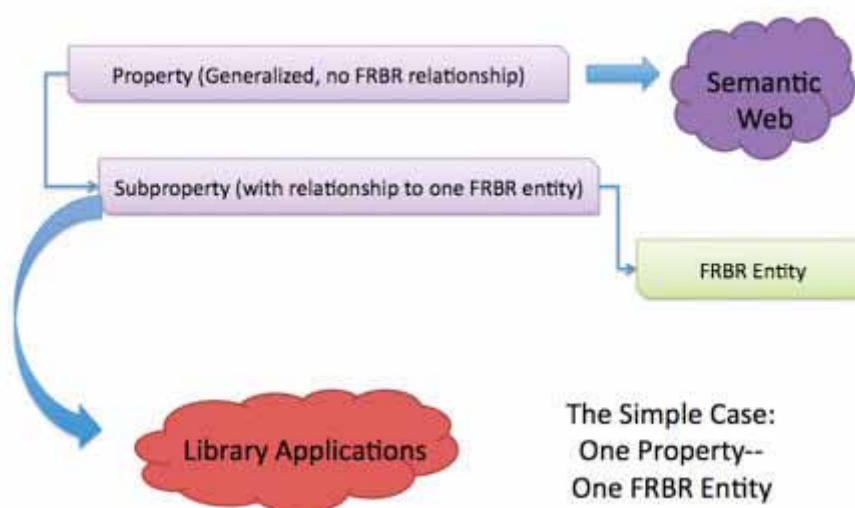
Onde não há entidade inerente FRBR atribuída ao nome de uma propriedade, duas propriedades foram criadas para os elementos RDA, conforme pode ser visualizada na figura 16:

1. Uma propriedade geral com nenhuma atribuição explícita de FRBR;
2. Uma subpropriedade FRBR limitada

A propriedade geral não carrega associações específicas com uma entidade FRBR e pode ser usada por qualquer aplicação que determina que seja útil para seu contexto, se a aplicação é ou não baseada em FRBR.

Estas propriedades gerais são totalmente compatíveis com a *Web Semântica* e não apenas específicas para aplicações em bibliotecas digitais. Além disso, a propriedade geral pode ser usada para estender RDA, associando-se a propriedade em um nível de aplicação com uma entidade FRBR diferente, que não seja a escolhida pelos desenvolvedores RDA.

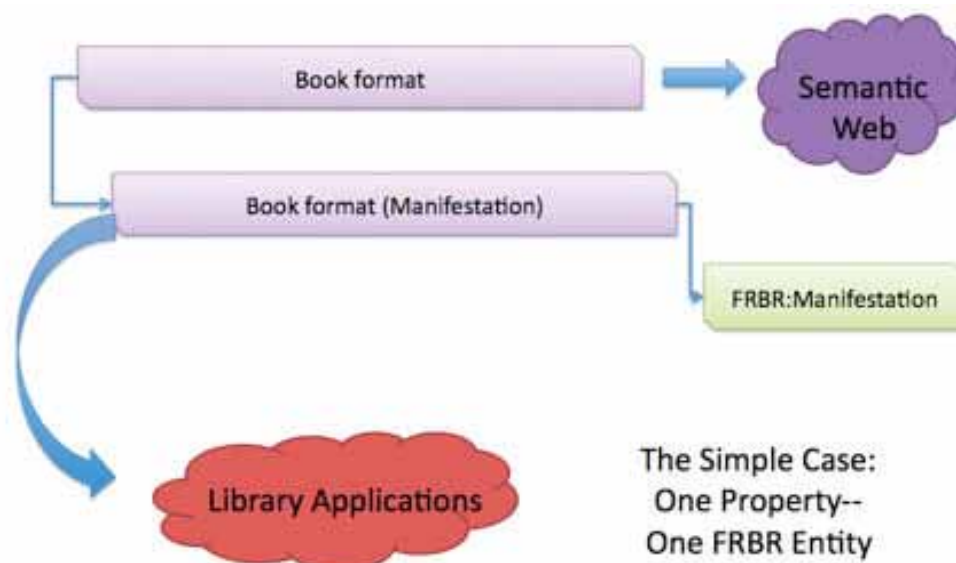
FIGURA 16: Criação de propriedades RDA para entidades FRBR.



Fonte: Hillmann et al. (2010, p. 7)

Como um exemplo, de acordo com a figura 17, a propriedade “*Book format*”, a qual é definida nas orientações de RDA com um relacionamento FRBR único, aparece no registro como uma propriedade geral com nenhum relacionamento FRBR definido, e novamente usando a definição “*Format book*” de RDA com um relacionamento explicitado para Manifestação FRBR. A adição do relacionamento FRBR cria uma propriedade com uma definição mais restrita e, portanto, é definida como uma subpropriedade mais geral, FRBR-less property.

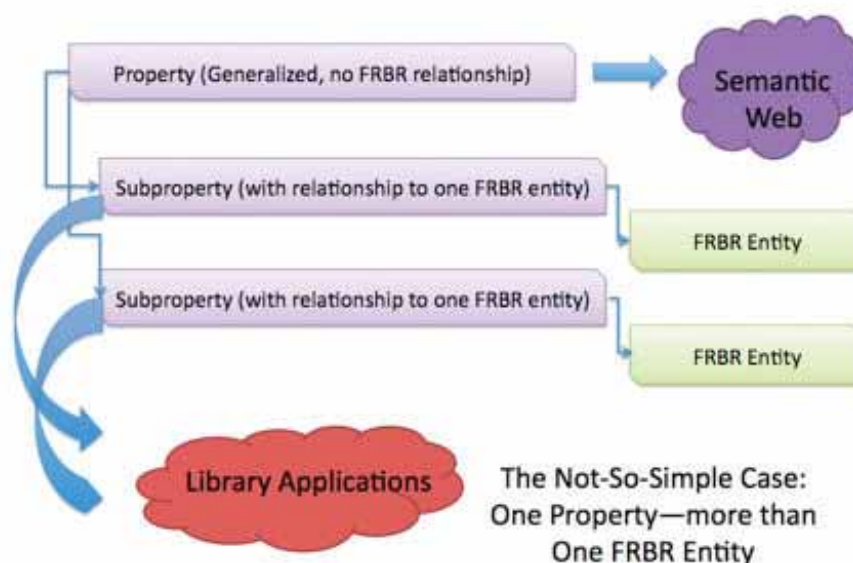
FIGURA17: Propriedade livro em RDA.



Fonte: Hillmann et al. (2010, p. 7)

A declaração de propriedades generalizadas que tem subpropriedade com relacionamentos específicos para entidades FRBR também acomoda o caso mais complexo de elementos que estão associados com mais de uma entidade FRBR em RDA. Esta abordagem, de acordo com a figura 18, fornece uma harmonia entre a necessidade de torná-la possível para os usuários a cumprir rigorosamente as especificações de RDA, e também permitindo uma flexibilidade importante para outros membros de comunidades sem a necessidade ou o desejo de obedecer a esse nível. (HILLMANN et al., 2010).

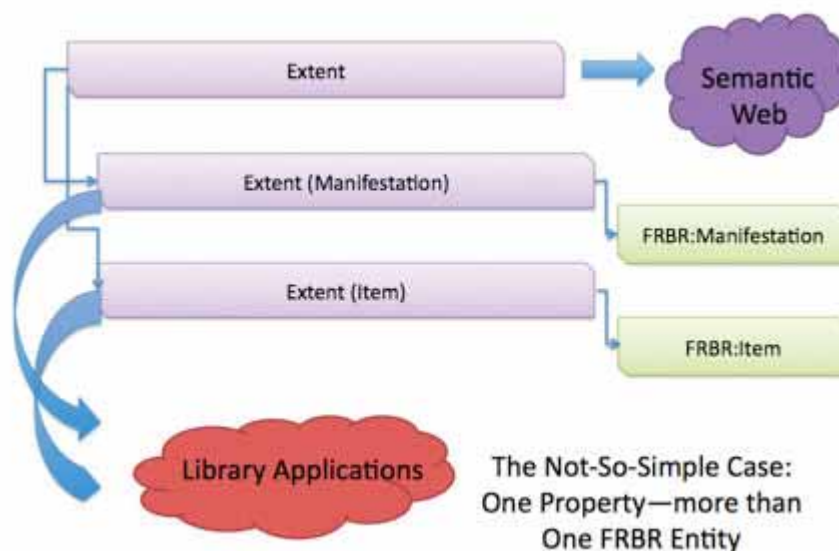
FIGURA 18: Propriedades e subpropriedade RDA e FRBR.



Fonte: Hillmann et al. (2010, p. 7)

Na figura 19, os relacionamentos com mais de uma entidade FRBR podem ser acomodados, tal como é o caso com o elemento RDA “*Extent*”.

FIGURA 19: Criação do elemento Extent.



Fonte: Hillmann et al. (2010, p. 8)

Os Diagramas de Entidade Relacionamento (DERs) criados para uso com a versão RDA *online* foram a principal fonte para as definições das propriedades RDA e seus relacionamentos para as entidades FRBR.

Porque o produto RDA *online* requer um Diagrama de Entidade Relacionamento separado para cada entidade, e não está claro a partir daquele DER que os elementos RDA estão relacionados a múltiplas entidades. Quando os elementos e os sub-elementos aparecem várias vezes nos DERs, não há indicação de que eles podem ser repetidos em outros lugares e se eles carregam a mesma definição etc. As propriedades RDA em RDF realizam esses relacionamentos explícitos através dos relacionamentos das propriedades/subpropriedade. (HILLMANN et al., 2010, p. 7, tradução nossa).

A disponibilidade de relacionamentos entre entidades descritos em FRBR é algo novo para as regras dos catálogos de bibliotecas, e é sem dúvida o aspecto mais inovador de RDA. Alguns dos relacionamentos são completamente novos para o processo de catalogação em bibliotecas, tais como os relacionamentos entre Obras. Outros relacionamentos, em particular de Pessoas para Obras e Expressões, têm sido expressados tradicionalmente em metadados de

bibliotecas como atributos de pessoas responsáveis (geralmente o autor) por um determinado recurso bibliográfico. (HILLMANN et al., 2010).

Em MARC 21, regras são atributos adicionados ao nome, tal como o campo/etiqueta/metadado 700 \$a Smith, John, \$e ed. Isso indica que uma pessoa (campo 700) com o nome “Smith, John” (\$a) é o (\$e) editor do recurso bibliográfico. Em RDF, as regras seriam um relacionamento, da mesma forma que uma pessoa coletiva (entidade) tem o relacionamento “editor” para o recurso informacional. (HILLMANN et al., 2010, p. 7, tradução nossa).

O Apêndice I do relatório final de RDA⁴² cobre as regras de RDA (relacionamentos entre as entidades dos Grupos 1 e 2 dos FRBR), tratando estes relacionamentos diferentemente dos outros estabelecidos em RDA, possivelmente para melhor suportar o mapeamento para MARC 21.

Com base nas instruções anteriores do *Joint Steering Committee for Development of RDA* (JSC), particularmente os comentários do editor sobre as regras de sintaxe RDF e do *DCMI Abstract Model*, os seguintes fatores foram usados na compilação das regras listadas no projeto:

- Não deve haver suposições a respeito de como as designações devem ser codificadas (por exemplo, representado como “add-ons” em termos relacionados em MARC 21 ou algum outro método), porque os dados em RDA devem funcionar independentes do formato, do meio ou do sistema utilizado para armazenar ou comunicar os dados.
- Para satisfazer as especificações RDF e ao *DCMI Abstract Model*, cada designação das regras deve funcionar como um elemento subtipo de um (e apenas um) dos elementos de nível superior definidos no Capítulo 6 (por exemplo, criador, colaborador etc.). Se uma dada regra é adequada para mais de um elemento, diferentes designações devem ser feitas. (HILLMANN et al. 2010; TILLET, 2008, tradução nossa).

A equipe de vocabulário DCMI baseou sua abordagem nos relacionamentos em geral sobre os trabalhos relevantes sobre as regras criadas em 2005, quando o *Dublin Core*

⁴² Esta diferença de abordagem foi reconhecida pelos autores do RDA no próprio documento (Apêndice I). Copyright © 2008, *American Library Association, Canadian Library Association, and The Chartered Institute of Library and Information Professionals*.

Metadata Initiative e a *Library of Congress* colaboraram para construir uma representação formal de *MARC Relators*⁴³, permitindo que estes termos sejam usados com Dublin Core.

Este trabalho anterior forneceu um modelo útil e totalmente vetado para a estratégia de desenvolvimento básica para os termos nas regras do Apêndice I de RDA, bem como outros relacionamentos nos Apêndices J e K. Todos os relacionamentos são, portanto, definidos como propriedades no mesmo nível que os elementos descritivos, ao invés de atributos, como estão em MARC 21. Esta é uma das primeiras formas para garantir que os elementos e os vocabulários são compatíveis para uso em RDF e com dados vinculados, bem como em XML. (HILLMANN et al., 2010).

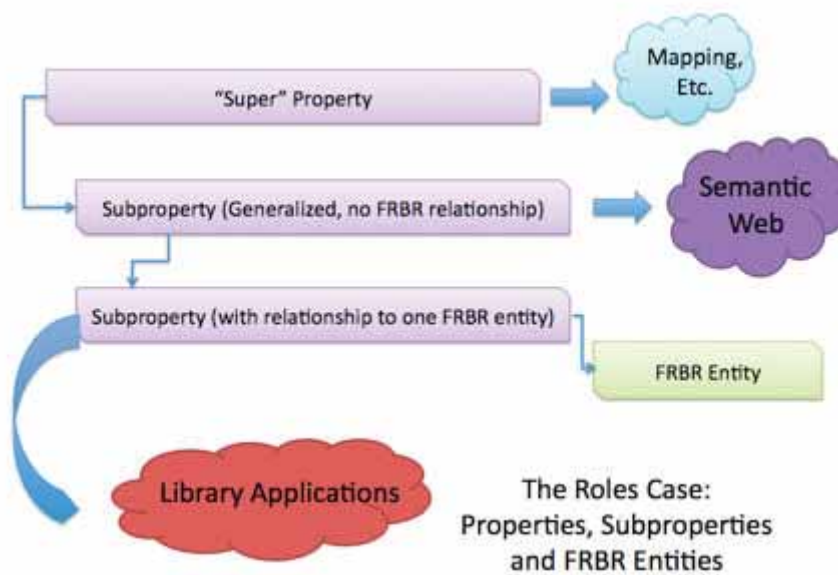
Todavia, algumas regras dos termos DERs foram estabelecidas através das associações com as propriedades “*Creator*” (Criador) e “*Contributor*” (Colaborador) do padrão de metadados Dublin Core, embora RDA tenha redefinido algumas propriedades DC. Todos os termos associados a um “*Creator*” foram também relacionados à entidade Obra dos FRBR, mas nem todas as regras relacionadas à entidade Obra dos FRBR foram relacionadas ao “*Creator*”. (HILLMANN et al., 2010).

Segundo Hillmann et al. (2010) um padrão semelhante ocorreu com “*Expression*” (Expressão) e “*Contributor*” (Colaborador). Havia também outras categorizações, para instância, as regras usadas com a “*Work*” (Obra) que não foram relacionadas ao “*Creator*” apareceram sob a categoria “*Other person, family, ou “corporate body”*” (outra pessoa, família ou entidade) associada à *Work* (Obra).

Embora não tenha sido difícil ver onde a caracterização *Creator/Contributor* deve ser útil no mapeamento dessas regras para o formato de intercâmbio bibliográfico MARC 21, campos 1XX/7XX - (a codificação das *tags* para a entrada principal e o acréscimo de pontos de acesso para nome pessoal – entrada principal nos tradicionais catálogos em fichas), a utilidade das caracterizações adicionais em um elemento de vocabulário baseado em RDF é menos claro, assim como aquelas “outras” caracterizações não foram incluídas, como pode ser visto na figura 20. (HILLMANN et al., 2010).

⁴³ Maiores informações sobre as representações formais de *MARC Relators* podem ser encontradas no *site*. Disponível em: <http://dublincore.org/documents/usageguide/appendix_roles.shtml>. Acesso em: 29 nov. 2011.

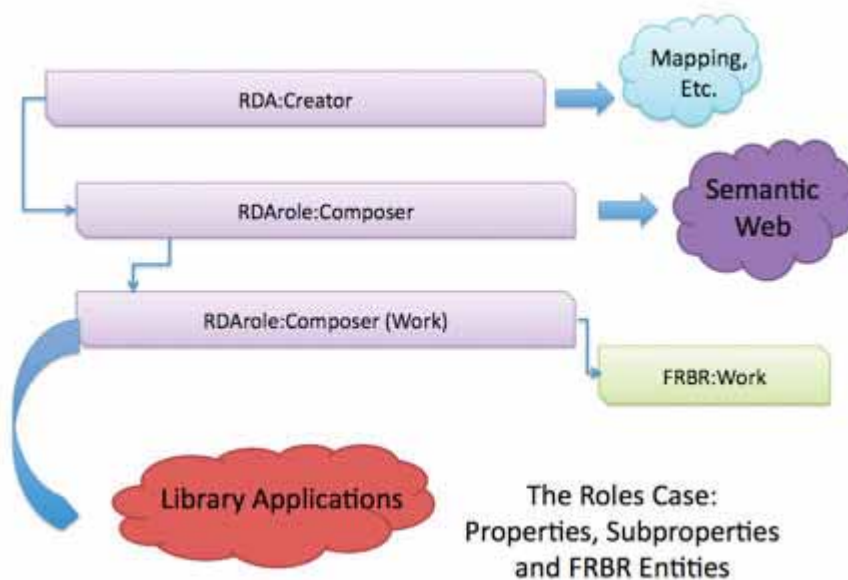
FIGURA 20: Caracterizações não incluídas no modelo RDA e RDF.



Fonte: Hillmann et al. (2010, p. 9)

Como exemplo, a função de "Composer" é listada como associada ao "Creator" e estaria relacionada, conforme mostra a figura 21.

FIGURA 21: Atributo Creator.



Fonte: Hillmann et al. (2010, p. 9)

Relacionamentos entre *Works* (Obra), *Expressions* (Expressão), *Manifestations* (Manifestação), *Items* (Item) (como consta na versão revisada do RDA, no Apêndice J, embora não nos DER) incluem o nome da entidade FRBR em cada nome da propriedade. De acordo com a figura 21, cada um dos relacionamentos (Apêndice J) traz definições distintas, geralmente refletindo um relacionamento específico da entidade FRBR. Nos vocabulários RDA, os relacionamentos são estruturados, como as outras propriedades RDA, com propriedades generalizadas de nomes mais específicos e propriedades limitadas de FRBR, para realizar um resultado similar. (HILLMANN et al., 2010).

Porque os relacionamentos frequentemente contêm três níveis de hierarquia (e essas podem variar entre as entidades), propriedades generalizadas foram criadas “rotineiramente” em cada nível para permitir níveis hierárquicos adequados para ser mantidos claramente, ambos com propriedades generalizadas e limitadas. Por enquanto, as propriedades generalizadas (*Generalized proprieties*) não deverão conter definições, como as que aparecem na subpropriedade (*Subproperty*), usada na entidade FRBR como uma parte da definição, fazendo a construção de definições gerais como uma tarefa posterior. (HILLMANN et al., 2010).

Segundo Hillman et al. (2010), no caso especial da propriedade “Name” (Nome), a descrição é relacionada especificamente a duas entidades do Grupo 2 de FRBR, “*Person*” (Pessoa) e “*Corporate Body*” (Entidade) e à “*Family*” (Família) de FRAD. Depois de algumas discussões a decisão tomada foi registrar a classe FRBRoo “*Agent*” no FRBR em vocabulário RDA e fazendo “*Person*”, “*Corporate Body*” e “*Family*” subclasses de “*Agent*”, permitindo o relacionamento a ser feito no nível “*Agent*” sem a duplicação desnecessária ao nível da subclasse. Para várias das propriedades que aparecem em cada DERs do Grupo 2, por exemplo, “*Note*” (Notas), “*Source consulted*” (Fonte consultada) e “*Status of identification*” (Status de identificação), a classe “*Agent*” é usada para fazer o relacionamento FRBR, por razões similares de eficiência, as propriedades do Grupo 3 são tratadas de forma semelhante. Algumas dessas “eficientes” decisões foram motivadas pela visão de que estas áreas do RDA não estão totalmente formuladas, e com o desenvolvimento e consolidação futuros, poderão ser revistas.

Para Hillmann et al. (2010) o RDA configura declarações, tais como *Publication*, *Distribution*, *Manufacture* e *Production*, e outros elementos de grupos pré-coordenados (*pre-*

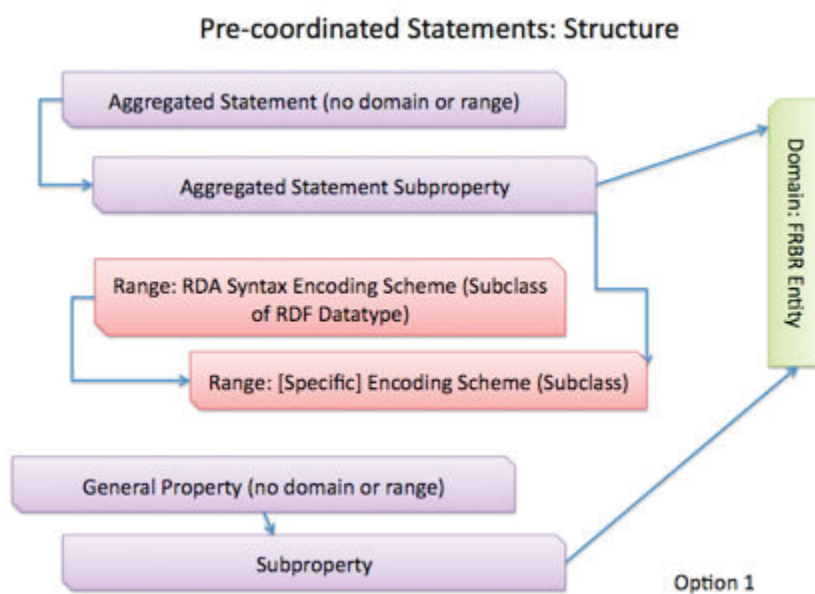
coordinated groupings) muito parecido ao trabalho realizado há muito tempo pelos bibliotecários na confecção de fichas catalográficas e o desenvolvimento de catálogos bibliográficos. Como exemplo, a agregação pré-coordenada de “*Place of publication*” (Lugar de publicação), “*Publisher’s name*” (Casa publicadora) e “*Publication date*” (Data de publicação) fornece o valor principalmente como uma opção, quando “*strings*” de texto são utilizados, como provavelmente será o caso quando os registros MARC são inicialmente transformados em RDA. Outros elementos, tais como Títulos uniformes (*Uniform titles*) são vistos principalmente como identificação, e não estão incluídos como declarações agregadas específicas nos vocabulários. Essas construções sintáticas oferecem um grande valor como formalizações de alguns dos conteúdos baseados nos *strings* de alguns campos do formato MARC 21. Eles são muito úteis para definir os dados que aparecem em um cartão físico e/ou em um catálogo de fichas. Além disso, eles fornecem uma estrutura “semi-semântica” em uma representação XML de dados RDA e MARC 21, conforme apresentado em seguida:

```
<Publication statement>
  <Place of publication> Austin, TX </Place of publication>
  <Publisher's name>The University of Texas at Austin, College of Liberal Arts </Publisher's
name>
  <Date of publication> [2001]- </Date of publication>
</Publication statement>
```

260 \$a Austin, TX : \$b The University of Texas at Austin, College of Liberal Arts, \$c [2001]-

Segundo Hillmann et al. (2010) no contexto RDF, dada a sua natureza composta, as declarações não contribuem muito com seus valores para a descrição de recursos informacionais. Este uso da pré-coordenação como uma parte inerente à definição de propriedade teve o resultado não intencional de limitar a capacidade para outros não interessados nessas declarações pré-coordenadas, para usar as propriedades RDA fora da aplicação em bibliotecas. Esse interesse foi validado quando a equipe composta por dois desenvolvedores da *Web Semântica* disse que queria usar “*Place of publication*”, mas não poderiam, porque em seus passos iniciais, eles tinham relacionado seu uso muito restritamente para agregar estruturas de declaração de Publicação. A última solução foi fornecer, tanto a capacidade para usar os elementos dentro das agregações, bem como separadamente, em parte pelo conceito explícito das agregações por elas mesmas como “Esquemas de Codificação da Sintaxe”, definido no *Dublin Core Abstract Model*.

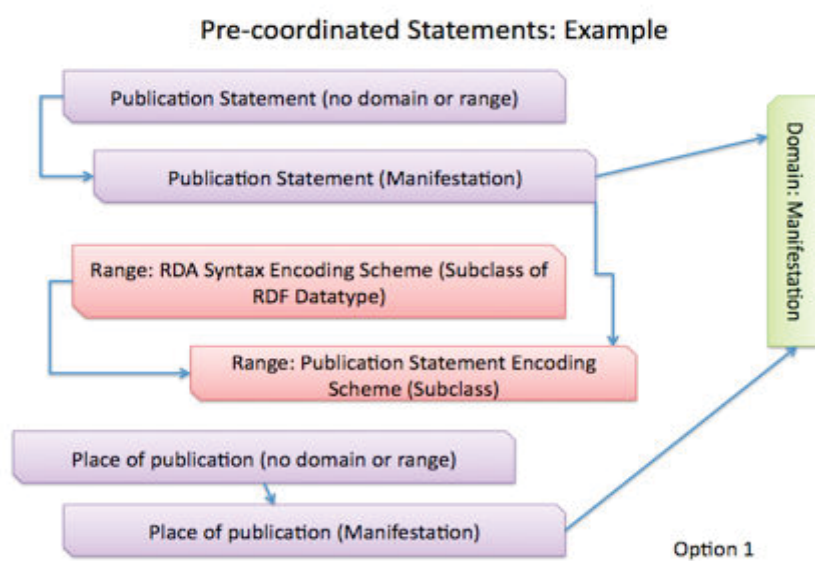
FIGURA 22: Estrutura de declarações pré-coordenadas.



Fonte: Hillmann et al. (2010, p. 10).

Como exemplo, a figura 23 apresenta uma declaração de Publicação estruturada da seguinte forma:

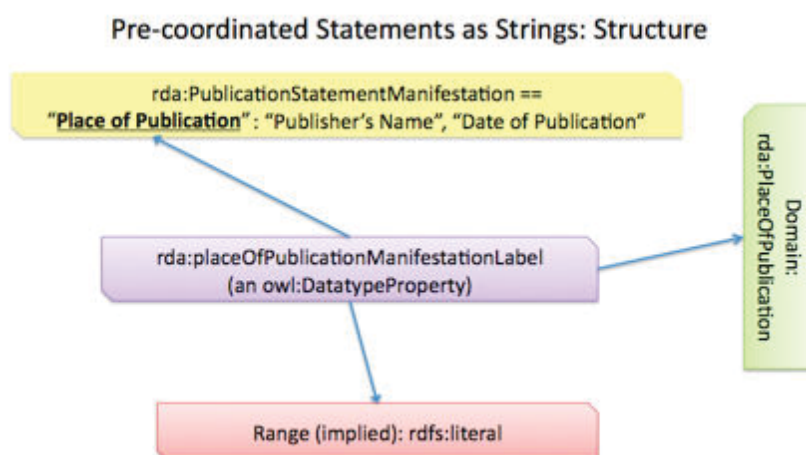
FIGURA 23: Declaração de Publicação.



Fonte: Hillmann et al. (2010, p. 10)

Hillmann et al. (2010) afirmam que uma parte significativa da mudança de expressar *Syntax Encoding Scheme (SES)* é a falta de uma forma padrão de descrever a sintaxe de um SES no modelo de dados RDF – nós podemos dizer que um **rda:publicationStatementManifestation** tem um tipo de dado de **rda:PublicationStatementEncodingScheme** mas não pode declarar facilmente a composição e a ordenação das *strings* requeridas. Além disso, DCMI não tem sido explícito sobre associações ao *DC Abstract Model (DCAM)* com a linguagem OWL, embora seja explícita sobre algumas definições com muitas restrições semelhantes. Por exemplo, para uma ontologia ser compatível com OWL-DL, deve declarar explicitamente se o objeto da tripla é um recurso (**owl:objectProperty**) ou um *literal* (**owl:datatypeProperty**). A especificação da linguagem OWL permite ambiguidade, mas ao custo de compatibilidade reduzida com “raciocinadores” padrão. O simples DCAM que você deve escolher e seu esquema não devem ser ambíguos. Uma possível solução é ilustrada, conforme a figura 24.

FIGURA 24: Declaração pré-coordenada como *string*: estrutura



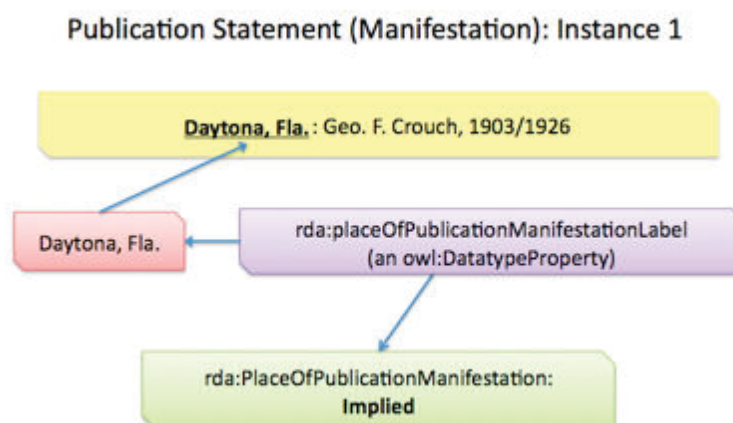
Fonte: Hillmann et al. (2010, p. 10)

Na figura 24, como é apresentada por Hillmann et al. (2010) foi requerido “*Place of Publication*” para suportar o maior número de dados bibliográficos possível, então definimos como as propriedades do RDF *Schema* são melhores do que as propriedades do OWL e não definem nenhum domínio ou extensão – e isso se torna efetivamente parte de uma ontologia de alto nível, como uma ontologia que pretende ser localizada e abrangente sendo mais usual. Os autores elucidam a necessidade que OWL-DL e DCAM sejam compatíveis, embora se

tenha criado uma específica manifestação com a subpropriedade **owl:objectProperty: 'Place of Publication (Manifestation)'** com o domínio de **rda:Manifestation** e uma extensão determinada de **rdf:resource**.

Isso permite instanciar dados para identificar o lugar da publicação usando URI em vez de um *literal*. Mas isso cria um problema para nosso SES, desde a definição da declaração de Publicação SES consiste de uma lista ordenada de *strings*. Nós também criamos um problema para aqueles que necessitam, porque os dados legados, para identificar o *Place of Publication* usando um *literal*, mas ainda querendo que sua ontologia seja compatível com OWL-DL. (HILLMANN et al., 2010).

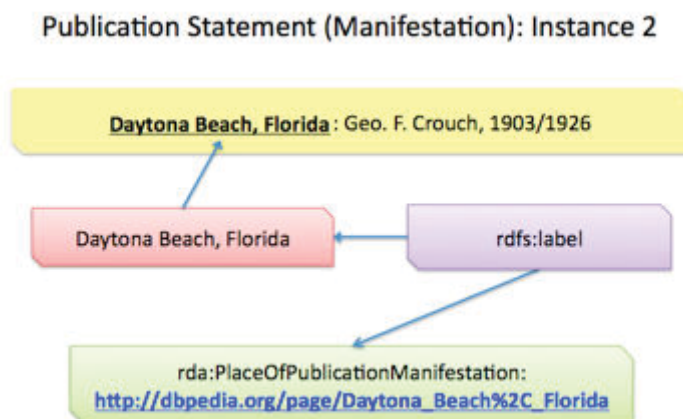
FIGURA 25: Instância 1: Lugar de publicação.



Fonte: Hillmann et al. (2010, p. 10)

Para solucionar o problema mencionado, Hillmann et al. (2010) disseram que foi criado um novo **owl:datatypeProperty** — **rda:PlaceOfPublicationManifestationLabel**, que tem um **rdfs:domain of rda:PlaceOfPublication** e implica em **rdfs:range of rdfs:Literal**. Esta propriedade, quando aplicada a um recurso, implicará a existência de um recurso **rda:PlaceOfPublication** não identificado, ou, se provado como uma propriedade de um recurso **rda:PlaceOfPublication**, ele pode então ser utilizado para fornecer um *string literal* para *Publication Statement SES*. (FIGURA 26).

FIGURA 26: Instância 2: Lugar de publicação.



Fonte: Hillmann et al. (2010, p. 10)

Os registros de metadados tem sido um tópico de discussão há mais de uma década. Com o aumento exponencial de informações, conseqüentemente ocorre o crescimento dos formatos e/ou padrões de metadados, e nesse contexto, a idéia de reaproveitar e combinar os diversos elementos da descrição é determinante.

Um registro público fornece informações sobre os padrões de metadados em uma estrutura legível por máquinas, capazes de realizarem a integração em aplicações específicas. A utilização de elementos de dados registrados aumenta a consistência do uso dos mesmos metadados, sobre uma variedade de serviços, uma vez que todas as aplicações são construídas sob as mesmas definições determinadas pela máquina. (HILLMANN et al., 2010).

Segundo Hillmann (2010), um registro de metadados eficiente facilita a declaração, o gerenciamento e a descoberta de esquemas de metadados legíveis por computador, perfis de aplicação e vocabulários controlados melhor estruturados. Com o registro de metadados a capacidade de interoperabilidade aumenta, assim como a precisão de qualquer compartilhamento de elemento de dados e vocabulários. Dessa maneira, os registros fornecem o significado para uma rede de informações global, com uma interação muito maior entre os serviços de informação.

De acordo com Hillmann et al. (2010) um aspecto-chave de um registro que deve ser considerado é que ele pode fornecer um único identificador (URI) para cada elemento de dados e para cada membro de um vocabulário.

Hillmann et al. (2010) elucida que um objetivo importante do JSC RDA é destacado, onde as orientações podem ser utilizadas além da comunidade bibliotecária, registrando elementos e conceitos que estão disponíveis para uso em aplicações em ambientes digitais com XML ou RDF, potencializando a codificação futura da descrição dos recursos informacionais. Além disso, o RDA concomitantemente com XML e RDF apoiam a construção de vocabulário e dos relacionamentos para a comunidade da *Web Semântica*.

O desenvolvimento de RDA implica para as bibliotecas condições de avançar rapidamente para um serviço de intercâmbio e de interoperabilidade mais amplo, reutilizando os dados bibliográficos de outros ambientes informacionais digitais. Os elementos e vocabulários RDA fornecem a base para a migração do uso exclusivo do formato MARC, que é relevante e significativo no domínio bibliográfico, transpondo fronteiras para outras comunidades compreenderem, interpretarem e utilizarem as informações mais amplamente.

Discussões recentes sobre o futuro da biblioteca apontam para a realidade que os grandes consumidores de metadados bibliográficos como *Amazon* e *Google Books*, têm usado dados em formato MARC de forma que revelam uma certa falta de compreensão dos metadados de uma biblioteca tradicional/convencional. (HILLMANN et al., 2010). Tal fato leva mais uma vez a reforçar a tese de que a compreensão e a aplicação do formato MARC 21 é ajustada somente por um catalogador especialista (bibliotecário), pois este detém as habilidades e as ferramentas necessárias na construção e na codificação dos metadados, apoiado em regras e esquemas de descrição que norteiam a representação adequada, padronizada e unívoca dos recursos bibliográficos no ambiente digital.

Destarte, o capítulo 5 pretende apresentar de que forma as metodologias da Catalogação Descritiva e uma ontologia de descrição bibliográfica explicitada nos códigos e formatos de metadados podem redesenhar novas abordagens sobre a construção de ambientes informacionais digitais contemporâneos, sobretudo, as tecnologias que estão sendo utilizadas e que permeiam o delineamento de tais ambiências, na garantia de propiciar uma recuperação da informação de maneira mais efetiva.



CAPÍTULO 5



**CATALOGAÇÃO DESCRITIVA NO SÉCULO XXI:
PROPOSTA DE UM MODELO PARA A DESCRIÇÃO
BIBLIOGRÁFICA SEMÂNTICA EM AMBIENTES
INFORMACIONAIS DIGITAIS INTEROPERÁVEIS**

5 CATALOGAÇÃO DESCRITIVA NO SÉCULO XXI: proposta de um modelo para a descrição bibliográfica semântica em ambientes informacionais digitais interoperáveis

A Catalogação tradicional abriu caminhos para o catalogador reordenar o que já havia sido feito e caminhar para a construção de catálogos estruturados e interoperáveis. *Plácida L.V.A.C. Santos*

Conforme pode ser verificado, muitos esforços têm acontecido na tentativa de se estabelecer padrões internacionais que garantam e promovam a interoperabilidade em ambientes informacionais digitais.

A preocupação da comunidade biblioteconômica advinda com as novas regras para a descrição bibliográfica aplicadas no ambiente digital, levam os profissionais a repensarem o seu fazer profissional, na tentativa de caminharem em consonância com as novas transformações trazidas no bojo da recontextualização tecnológica das bibliotecas.

Dessa forma, pretende-se apontar e refletir sobre tais mudanças, numa proposta de um modelo de requisitos funcionais, para a modelagem de catálogos bibliográficos, pautados na lógica descritiva das metodologias da Catalogação Descritiva, nas ontologias de descrição bibliográfica explicitadas nas regras e códigos de catalogação e em padrões de metadados, no estabelecimento efetivo da interoperabilidade em ambientes informacionais digitais.

5.1 A Catalogação Descritiva e sua repercussão nos ambientes informacionais digitais

No cenário atual da comunidade da Biblioteconomia e da Ciência da Informação, há uma preocupação e um reconhecimento crescente de que haverá a necessidade de um sucessor do formato bibliográfico MARC 21, devido às novas transformações no domínio bibliográfico, permeada pelo uso intensivo das TIC. Segundo Coyle (2011) tais discussões

tendem a se concentrar principalmente nas questões estruturais, o novo formato será XML, ele fará uso de RDF e padrões de dados ligados⁴⁴?

O que essas questões não se referem é a tarefa muito mais complexa de traduzir a semântica dos dados bibliográficos para a construção de um novo padrão. De acordo com Thomale (2010) basta apenas uma pequena investigação dos dados codificados no formato MARC 21, para revelar que as *tags* e os subcampos em si, são insuficientes para definir os elementos dos dados reais realizados pelos registros de um catálogo bibliográfico. “O primeiro passo na transformação de MARC 21 para um outro formato é identificar quais são os elementos contidos num registro bibliográfico MARC 21, o que não parece tão simples”. (THOMALE, 2010, p. 3, tradução nossa).

O entusiasmo recém-desenvolvido para RDF se tornar a base para os dados bibliográficos de uma biblioteca tem ocorrido uma série de esforços e aplicações que convertam MARC 21 para RDF, mas nenhuma oficialmente ou recomendada pelos organismos internacionais da Catalogação Descritiva.

Dentre as iniciativas oficialmente aceitas da transformação dos dados em RDF encontram-se na *Library of Congress* (LC) padrões para a descrição de recursos, tais como o MODS (*Metadata Object Description Standard*). Outras propostas para a transformação dos dados de bibliotecas para RDF podem ser vistas em: ISBD⁴⁵ (*International Standard Bibliographic Description*) em RDF, FRBR⁴⁶ em RDF e RDA⁴⁷ em RDF.

Cada um desses esforços toma uma biblioteca padrão e usa RDF como sua tecnologia subjacente, criando um esquema de metadados completo que define cada elemento do padrão em RDF. O resultado é que se tem uma série de "silos" RDF, e cada elemento de dados definido como se pertencessem exclusivamente a esse padrão.

⁴⁴ O termo dados ligados (*Linked Data*) refere-se ao conjunto das melhores práticas para a publicação e a interligação de dados estruturados na *Web*. Essas melhores práticas foram introduzidas por Tim Berners-Lee em sua nota *Web Architecture* e tornaram-se conhecidas como princípios de dados ligados. (HEATH; BIZER, 2011, tradução nossa).

⁴⁵ Maiores informações no *site*: Disponível em: <<http://www.ifla.org/en/node/1795>>. Acesso em: 20 jan. 2012.

⁴⁶ Maiores informações no *site*: Disponível em: <<http://metadataregistry.org/schema/show/id/5.html>>. Acesso em: 20 jan. 2012.

⁴⁷ Maiores detalhes no *site*: Disponível em: <<http://metadataregistry.org/schema/show/id/1.html>>. Acesso em: 20 jan. 2012.

Existem quatro declarações diferentes, por exemplo, no elemento “lugar de publicação” em ISBD, RDA, FRBR e MODS, cada um com seu próprio URI e também há diferenças entre elas (por exemplo, RDA separa lugar de publicação, produção etc., enquanto ISBD não o faz), certamente deverão ter uma estrutura (vocabulário) comum para solucionar essas questões. (COYLE, 2012).

Uma possível solução seria se as diferentes instâncias do "lugar de publicação" pudessem ser tratadas como tendo um significado comum, tal que elementos de FRBR pudessem ser ligados a um elemento ISBD, mas o mesmo não ocorre.

A razão pela qual não ocorre se justifica, porque cada uma dessas restringe os elementos de forma única (individual) que define seu relacionamento com um contexto de dados particular (o que geralmente pensamos das estruturas de um registro). Os elementos não são independentes do contexto, e isso significa que cada um somente pode ser usado dentro desse contexto particular. Essa é a antítese do conceito de dados ligados, onde conjuntos de dados de fontes diversas compartilham elementos de metadados. E esta reutiliza os elementos que cria o "link" em dados ligados (vinculados). Para conseguir tal processo, os elementos de metadados precisam ser irrestritos por um contexto particular. (COYLE, 2012).

A ligação pode também ser alcançada através dos relacionamentos verticais, similares aos termos (mais amplos e mais restritos) num tesouro. Essa opção é menos direta, mas torna possível a "mistura" de conjuntos de dados que têm diferentes níveis de granularidade. No caso do "lugar de publicação" da ISBD, poderiam ser definidos mais amplamente para os três elementos de RDA que tratam esse separadamente. Coyle (2012) afirma que infelizmente isso não é possível devido à forma que ISBD e RDA foram definidos em RDF.

Coyle (2012) reforça que o resultado é que agora tem-se uma série de “silos” RDF, expressões dos dados em RDF que faltam a capacidade de cruzamento dos dados ligados, porque eles são obrigados a especificar estruturas de dados e pouco se ganha em termos de dados ligados sob a óptica bibliográfica. Não somente os esquemas RDF são incompatíveis com outros, ninguém será ligado a dados bibliográficos de comunidades fora de bibliotecas que publicarem seus dados na *Web*. O que significa a não ligação dos dados à Amazon, à *Wikipedia*, para citações em documentos.

Coyle (2012) diz que devido ao estágio inicial em que estamos no desenvolvimento de dados ligados para o ambiente das bibliotecas, têm-se duas opções em relação ao uso do RDF.

- Definir "super-elementos" que estejam acima dos registros de formatos e que não estão vinculados pelas restrições dos registros definidos em RDF. Neste caso, haveria um "lugar de publicação" geral que seria um "super" elemento correspondente a todos os lugares de publicação nos vários registros e seriam subordinados a um conceito geral de "lugar" utilizado amplamente. Para implementar a ligação, cada elemento de registro seria extrapolado para os seus "super-elementos".

- Definir primeiro os elementos de dados fora de qualquer formato de registro específico, e então usá-los em esquemas de registros. Neste caso haveria apenas uma instância do "lugar de publicação" e que seria utilizado nos vários registros bibliográficos, sempre que um elemento é necessário. O intercâmbio desses registros seria possível com a ligação dos dados de seus elementos componentes, e iria interagir com outros dados bibliográficos na *Web* utilizando os elementos RDF definidos e seus relacionamentos.

Coyle (2012) conclui ainda que precisa-se criar dados, não somente registros, e que necessita-se criar primeiramente os dados, em seguida, registros com os mesmos para aplicações de acordo com a necessidade de cada ambiente informacional. Esses registros irão operar internamente nos sistemas de bibliotecas, enquanto os dados têm o potencial para fazer conexões no espaço de dados ligados. Um esforço maior tem que acontecer em descobrir e definir os elementos de nossos dados e olhar para os diversos dados que deseja unir, no vasto universo de informações.

As bibliotecas têm seu foco nos registros bibliográficos que compõem os acervos institucionais, geralmente um documento complexo que atua como um substituto do catálogo, tal como um livro ou uma gravação de música. RDF, nesse contexto, não menciona nada sobre os registros, apenas diz que existem dados que representam coisas (recursos) e os relacionamentos entre essas coisas. O que é muitas vezes confundido é que qualquer coisa pode ser uma coisa em RDF, assim, o livro, o autor, a página, a palavra na página, qualquer uma, ou todas essas poderiam ser as coisas em seu universo. (COYLE, 2012).

As discussões acerca do futuro das bibliotecas digitais e sua configuração para a adoção de tecnologias da *Web Semântica* manifestam-se da necessidade que os ambientes e sistemas de informação criem uma estrutura de dados que aproveitem as potencialidades do RDF, permitam os relacionamentos de forma explicitada e promovam a interoperabilidade.

Vale destacar que os relacionamentos bibliográficos sempre existiram no interior dos catálogos bibliográficos entre os dados de um registro bibliográfico, por meio das regras e esquemas de catalogação, porém não explicitados aos sujeitos institucionais para o uso e (re) uso das informações nos ambientes de manipulação de dados bibliográficos e catalográficos.

5.2 Modelagem de dados para construção de ambientes digitais

Ao pensar na criação e no desenvolvimento de um modelo de dados, nos deparamos com as questões de granularidade e análise dos dados a serem catalogados. O que não é algo novo no domínio bibliográfico, onde se desenvolveu um modelo de dados implícito nas regras de descrição (AACR2) e nos formatos de intercâmbio de dados bibliográficos (MARC 21).

Atualmente com o desenvolvimento de novas regras e padrões para a modelagem dos ambientes informacionais, tais como FRBR, FRAD e RDA, reconhece-se que há uma tendência na estruturação e na definição dos dados a serem catalogados preparando-os para movê-los e torná-los compatíveis com a *Web Semântica*.

Nas camadas intangíveis dos dados (definida na representação e descrição dos recursos informacionais), há um aumento na estrutura e na granularidade dos dados. Yee (2009) aponta que mais estrutura e mais granularidade possibilitam apresentações mais sofisticadas aos usuários dos sistemas e aumentam a possibilidade de produção de dados interoperáveis.

Qualquer mudança ou mapeamento que foi contratado para criar dados interoperáveis produziria um menor denominador comum (os dados mais simples e menos granular), e uma vez interoperáveis, não seria possível sua recuperação na íntegra, devido sua perda. Dados com menos estrutura e menos granularidade poderiam ser mais fáceis e mais baratos para aplicá-los e ter o potencial mais simples para as comunidades envolvidas, mas que os

dados iriam limitar o grau em que a indexação e a apresentação seriam possíveis. (YEE, 2009, p. 59, tradução nossa).

Vamos tomar como exemplo um nome pessoal. Conforme as regras de catalogação (AACR2), demarca-se o sobrenome, do nome próprio, registrando primeiro o sobrenome, seguido por uma vírgula e depois o nome. Essa quantidade de granularidade pode representar muitas vezes um problema para o catalogador numa cultura desconhecida, que não domina necessariamente as regras. Mais granularidade pode ocasionar situações ambíguas para os sujeitos que estão coletando os dados. (YEE, 2009).

Yee (2009) comenta que se adicionarmos uma data de nascimento e de morte, seja qual for, as utilizamos juntas em um subcampo *\$d* sem qualquer codificação separada para indicar qual é o nascimento e qual é a morte (embora um ocasional "b" ou "d" nos dirá esse tipo de informação). Poderíamos fornecer mais granularidade para datas, mas tornaria o formato MARC 21 muito mais complexo e difícil de aprender.

Na representação do campo 100 (autor pessoal), por exemplo, a forma autorizada para descrever o conteúdo é definida da seguinte maneira:

100 1#\$a Adams, Henry, \$d1838-1918.

Nesse caso, o subcampo *\$d* (Datas associadas ao nome NR⁴⁸), 1838 corresponde à data de nascimento, enquanto 1918 à data de falecimento do autor.

De acordo com Yee (2009), granularidade e estrutura também podem ocorrer "tensão" uma com a outra. Mais granularidade pode conduzir a uma menor estrutura (ou mais complexidade para manter a estrutura junto com a granularidade). Na busca de maior granularidade de dados que se tem agora, (RDA tenta apoiar RDF na codificação XML), têm sido atomizados os dados para torná-los úteis aos computadores, mas isso não necessariamente tornará os dados mais compreensíveis aos humanos.

⁴⁸ (NR) Não Repetido, significa que só pode existir apenas um subcampo com apenas um código associado ao nome do responsável pela obra.

Para ser útil aos humanos, deve ser possível agrupar e organiza-los de forma significativa para a catalogação, a indexação e a apresentação dos mesmos. Os desenvolvedores do *Simple Knowledge Organization System* (SKOS) referem-se ao montante de informações não estruturadas (isto é, legível por humanos) na *Web*, rotulando *bits* de dados como relacionamentos semânticos dos registros em uma máquina acionável, de forma que não necessariamente fornece o tipo de estrutura necessária para tornar os dados legíveis por humanos e, portanto, úteis para as pessoas na *Web*. (YEE, 2009).

Para reforçar seu pensamento, Yee (2009, p. 59, tradução nossa) afirma que:

Quanto mais granular os dados, menos o catalogador pode construir ordem, sequenciamento e a ligação dos dados; a codificação deve ser cuidadosamente projetada para permitir a ordem, o sequenciamento e a ligação dos dados desejados, para que a catalogação, a indexação e a apresentação sejam possíveis, o que poderia ser chamado de uma codificação dos dados ainda mais complexa.

No que tange à estrutura de dados, Yee (2009) diz que existem vários significados atrelados ao termo, conforme pode ser observado.

- Estrutura é um objeto de um registro (estrutura de documento), por exemplo, *Elings* e *Waibel* referem-se a "campos de dados... também referidos como elementos... que estão organizados em um registro por uma estrutura de dados".
- Estrutura é a camada de comunicação, ao contrário da camada de apresentação ou designação de conteúdo.
- Estrutura é o registro, o campo e o subcampo.
- Estrutura é a ligação de *bits* de dados em conjunto, na forma de vários tipos de relacionamentos.
- Estrutura é a apresentação dos dados de maneira estruturada, ordenada e sequenciada para facilitar a compreensão humana.
- Estrutura de dados é a forma de armazenamento dos dados em um computador para que ele possa ser usado eficientemente (isto é, como os programas de computadores usam o termo).

Conforme as definições apresentadas por Yee (2009) para essa tese chamar-se-á de estrutura de dados, a camada intangível de instanciamento dos dados bibliográficos

modelados para a representação e a descrição, tal como os formatos e/ou padrões de metadados, de forma a permitir a interoperabilidade dos ambientes informacionais digitais, por agentes humanos e não humanos, garantindo interfaces mais acessíveis aos usuários para posterior recuperação, uso e (re) uso dos recursos informacionais.

Quando se menciona estrutura de dados, no domínio bibliográfico, pensa-se no modelo conceitual de dados, estabelecido pelos Requisitos Funcionais para Registos Bibliográficos (FRBR).

FRBR faz uso de um modelo entidade-relacionamento, o qual consiste em dois principais conceitos: "coisas" e relacionamentos. FRBR define 10 categorias "coisas", as quais são denominadas entidades: Obra, Expressão, Manifestação, Item, Pessoa, Entidade coletiva, Conceito, Objeto, Evento e Lugar.

As entidades podem ser compreendidas, por exemplo, como uma obra, um texto, um livro etc. Os atributos correspondem às características dos dados relacionados à entidade e servem para diferenciar o conteúdo intelectual ou artístico. Os relacionamentos descrevem as ligações entre uma entidade e outra, na facilitação de manuseio do recurso informacional pelo usuário em um sistema. (MORENO, 2006).

No que se refere à entidade, nos FRBR vale dizer que esta é compreendida por 3 (três) grupos:

- **Grupo 1:** entidades que são produto de trabalho intelectual ou artístico,
- **Grupo 2:** entidades que são responsáveis pelo conteúdo intelectual, guarda ou disseminação das entidades do primeiro grupo e,
- **Grupo 3:** entidades que são ou podem ser assunto das entidades.

Moreno (2006, p. 35, grifo do autor) apresenta as principais terminologias no modelo conceitual de dados FRBR:

Obra é uma entidade abstrata, uma criação intelectual ou artística distinta. A entidade **Expressão** é a realização intelectual ou artística específica que assume uma obra ao ser realizada, excluindo-se aí aspectos de alteração da forma física. Uma **Manifestação** é a materialização de uma expressão de uma obra, ou seja, seu suporte físico, que podem ser livros, periódicos, kits

multimídia, filmes, etc., que é representada pelo **Item**, um único exemplar de uma manifestação. As duas últimas entidades refletem a forma física, são entidades concretas, enquanto as duas primeiras refletem o conteúdo intelectual ou artístico.

A contextualização e a exemplificação dos FRBR podem ser interpretadas da seguinte maneira. Tomemos como exemplo, um romance literário (Obra), onde se tem o texto original e o mesmo fora traduzido ou tivera alguma modificação, tal como uma edição ilustrada (Expressão); as formas em que a obra está disponível podem ser encontradas/visualizadas em um formato impresso (manual) ou num formato eletrônico/digital (Manifestação); e no momento em que a obra está disponibilizada na estante de uma biblioteca, ou seja, os exemplares relativos àquela obra são denominados **Itens** (“materialização” do recurso bibliográfico).

Os FRBR constituem em uma “nova”⁴⁹ abordagem para a Catalogação Descritiva contemporaneamente nos seus moldes convencionais, no sentido de propiciar uma recuperação mais efetiva e intuitiva dos itens bibliográficos, agindo como um bibliotecário de referência, ou seja, relaciona todos os materiais atrelados ao termo da busca, trazendo-os de uma só vez em uma única interface. Por exemplo, se um determinado autor além de livros, possui outras manifestações, tais como discos, CDs e DVDs, o sistema/agente permitirá através do modelo de relacionamento FRBR, no momento da busca realizada pelo usuário, relacionar todas essas manifestações e recuperá-las uma única vez, apresentando-as aos usuários.

Os bibliotecários que se acostumarem com os Requisitos Funcionais para Registros Bibliográficos - FRBR - provavelmente não encontrarão muita dificuldade na transição do modelo conceitual para o modelo RDF. É importante nesse momento destacar as principais terminologias empregadas pelos FRBR, RDF e RDFS, de acordo com o Quadro 11.





⁴⁹ “Nova” no sentido de explicitar os relacionamentos entre registros bibliográficos, uma prática que já acontece, por exemplo, entre os elementos do formato de intercâmbio MARC 21, mas de forma não visível aos usuários dos catálogos *on-line*.

Quadro 11: Diferenças terminológicas nos modelos de dados.

FRBR	RDF	RDFS
Entidade	Assunto	Classe
Atributo	Objeto	Propriedade
Relacionamento	Predicado	Relacionamento/Relação Semântica

Fonte: Yee (2009, p. 64, tradução nossa).

Ainda nesse contexto Riley (2010) aprofunda o estudo comparativo terminológico do RDF no contexto da Biblioteconomia e da Ciência da Informação, especificando:

-  Assunto: em bibliotecas, o que um recurso informacional abrange em termos de conteúdo; em RDF – o que diz uma declaração sobre alguma coisa (recurso informacional).
-  Vocabulário: em bibliotecas, implica num certo tipo de vocabulário controlado (termos autorizados, estruturas hierárquicas, termos relacionados etc.); em RDF – definições muito mais flexíveis (inclui definições formais de classes e de propriedades de um recurso informacional).
-  Classe: em bibliotecas, um esquema de classificação (Classificação Decimal de Dewey – CDD; Classificação Decimal Universal – CDU etc.) indicando o tópico geral ou área do conhecimento abrangido pelo recurso informacional; em RDF - um tipo ou categoria pertencente a um objeto ou recurso informacional.
-  “Schema”: XML *Schema* define um conjunto de elementos destinados a ser usados juntos; RDF *Schema* define classes e propriedades destinadas a ser usadas em qualquer lugar, sozinhas ou em combinação.

A dificuldade em qualquer exercício de modelagem de dados, sobretudo, no domínio bibliográfico, encontra-se em decidir o que tratar como entidade ou classe e o que tratar como um atributo ou propriedade. FRBR decidiu criar uma classe chamada expressão para tratar qualquer mudança no conteúdo de uma obra.

Os FRBR em comparação ao modelo de dados RDF, encontra-se em harmonia, as entidades do FRBR são registradas como classes, enquanto as relações são registradas como propriedades.

FRBR em RDF⁵⁰ acrescenta apenas três classes. Duas delas (“*Endeavor*” e “Entidade Responsável”) são super conjuntos de classes FRBR. *Endeavor* é uma generalização que pode ser relacionada à obra, expressão, ou manifestação, ou seja, uma classe cujos membros são qualquer um dos produtos da atividade artística ou criativa. “Entidade Responsável” é um termo mais geral que pode se relacionar a uma entidade coletiva ou a uma pessoa. Essas classes especificam mais claramente informações sobre o conteúdo intelectual de um recurso, sem necessitar fornecer informações adicionais. “A terceira classe que é adicionada é o assunto. Todas as três incluem alguma instância do assunto em seus esquemas. FRBR trata claramente assunto como um relacionamento”. (DAVIS; NEWMAN, 2005; COYLE, 2012).

Com base na sustentação teórica apresentada nesse texto, até o presente momento, fica evidente que a condição dos ambientes informacionais digitais interoperar seus dados é a modelagem conceitual, esta definida e codificada por uma série de requisitos funcionais estabelecidos pelas arquiteturas de metadados, regras e esquemas de descrição bibliográfica e ontologias, que fornecerá ambientes melhor estruturados, na garantia de recuperação de informações mais efetiva aos usuários (humanos e não humanos).

5.3 Proposta de um modelo para a descrição bibliográfica semântica em ambiente digital

Percebe-se que a palavra de ordem na constituição de ambientes informacionais digitais, sobretudo no âmbito das bibliotecas digitais é a interoperabilidade. Para tanto, para

⁵⁰ As expressões do núcleo de conceitos e vocabulários FRBR em RDF podem ser melhor visualizados no *site*. Disponível em: <<http://vocab.org/frbr/core.html>>. Acesso em: 20 jan. 2012.

que a interoperabilidade aconteça de maneira eficiente, faz-se necessário um olhar mais acurado nas camadas estruturantes e delimitadoras dos catálogos bibliográficos digitais, ou seja, na representação e na descrição dos dados, a fim de potencializar as formas de busca e recuperação de informações.

Yee (2009) num estudo realizado sobre como os dados bibliográficos podem ser inseridos no contexto da *Web Semântica*, apresenta um comparativo de como os dados dos sistemas de bibliotecas são construídos atualmente e como poderão ser construídos no futuro (*Web Semântica e/ou Web de dados*).

Quadro 12: Possíveis realocações de funções atuais para a *Web Semântica*.

FUNÇÃO	ATUALMENTE	FUTURO
Dados de conteúdo, ou diretrizes de conteúdo (regras para fornecer dados em um elemento particular)	Definidos por AACR2 e MARC 21.	Definidos por RDA, RDF, RDFS, OWL e SKOS.
Elementos de dados	Definidos por AACR2, ISBD e MARC 21.	Definidos por RDA, RDF, RDFS, OWL e SKOS.
Valor dos dados	Definidos por arquivos de autoridade LC/NACO, LCSH, valores de dados codificador por MARC 21.	Definidos como ontologias usando RDF/RDFS/OWL/SKOS
Codificação ou rotulagem de elementos de dados para manipulação por máquinas	Definidos por ISO 2709, baseado em MARC 21.	Definidos por RDF/RDFS/XML
Estrutura de dados	Definidos por AACR2 e MARC 21.	Definidos por RDF, RDFS, OWL e SKOS.
Esquematização (restrição na estrutura e conteúdo)	MARC 21, MODS, DCMI <i>Abstract Model</i>	Definidos por RDF, RDFS, OWL e SKOS.
Codificação de dados sobre os relacionamentos das entidades	Realizada por sequências de valores de dados correspondentes (cabeçalhos encontrados em arquivos de autoridade LC/NACO, LCSH, ISSN's etc.).	Realizado por RDF, RDFS, OWL e SKOS na forma de links URI.
Regras de apresentação	<i>Softwares</i> baseados em ISBD, AACR2 etc.	Camadas de aplicação

Fonte: Yee (2009, p. 62, tradução nossa).

A estrutura de dados é tomada para representar o significado de um padrão de registro. Tradicionalmente um registro tem representado uma expressão (livros, por exemplo) ou uma manifestação (desde mecanismos de reprodução que se tornam mais sofisticados, permitindo uma explosão de reproduções do mesmo conteúdo em diferentes formatos e provenientes de diferentes distribuidores). (YEE, 2009).

Segundo Yee (2009, p. 62, tradução nossa),

RDA é um registro neutro; RDF permitiria o URI ser estabelecido por qualquer e para todos os níveis de FRBR, ou seja, haveria um URI particular para uma obra, uma expressão, uma manifestação e um item. Não está sendo usada a estrutura de dados no sentido que faz um programa de computador (como uma maneira de armazenar os dados em um computador para que ele possa ser usado de maneira eficiente).

Atualmente a codificação de dados sobre o relacionamento das entidades é realizado pela correspondência dos valores dos dados numa sequência de caracteres (cabeçalhos ou campos ligados usando ISSN etc.) que são definidos por arquivos de dados de autoridade de nomes do programa de catalogação cooperativa (LC/NACO⁵¹) e dados de autoridade de assuntos – *Library of Congress Subject Headings* (LCSH⁵²). No futuro, esta função pode ser realizada usando RDF para ligar o URI de um recurso informacional ao URI para um valor de dados. (YEE, 2009).

As regras de apresentação (*display*) são atualmente definidas pelas ISBD e AACR2, “mas amplamente ignoradas pelos sistemas, os quais frequentemente truncam registros bibliográficos arbitrariamente em níveis de apresentação (*displays*) e fornecimento de dados. RDA ‘abdica’ da responsabilidade, fornecendo apresentações (*display*) fora das regras de catalogação”. (YEE, 2009, p. 62, tradução nossa).

Para Yee (2009) o princípio geral na *Web* é de se “divorciar” dos dados de exibição e permitir que qualquer um apresente os dados da maneira como desejarem. As formas de apresentação dos dados (*display*) é o coração dos objetos (ou metas) da catalogação. O ponto é para apresentar ao usuário as obras de um autor, as edições de uma obra, ou as obras sobre

⁵¹ Maiores informações sobre o NACO (*Name Authority Cooperative Program*) estão no *site* da *Library of Congress*. Disponível em: <<http://www.loc.gov/catdir/pcc/naco/>>. Acesso em: 20 jan. 2012.

⁵² Mais detalhes do *Library of Congress Subject Headings* no *site*. Disponível em: <<http://www.loc.gov/aba/cataloging/subject/>>. Acesso em: 20 jan. 2012.

um assunto. Todas estes objetivos somente poderão ser alcançados aumentando a complexidade, ou seja, (*displays*) de alta qualidade construídos, a partir dos dados criados de acordo com o modelo de dados definido.

De acordo com Yee (2009, p. 63, tradução nossa),

O nosso interesse na Web Semântica deve concentrar-se sobre se ela irá ou não apoiar mais dados estruturados, bem como mais lógica na sua estruturação, para fornecer índices e displays melhores do que temos agora em nossas OPACs.

Pensando nas recomendações funcionais que garantem uma modelagem de ambientes informacionais mais efetiva, e nos requisitos propostos no quadro 10, pode-se fazer uma releitura e a determinação de funções em domínios específicos, para a construção de bibliotecas digitais semânticas e o aproveitamento de descrição bibliográfica para a *Web Semântica*.

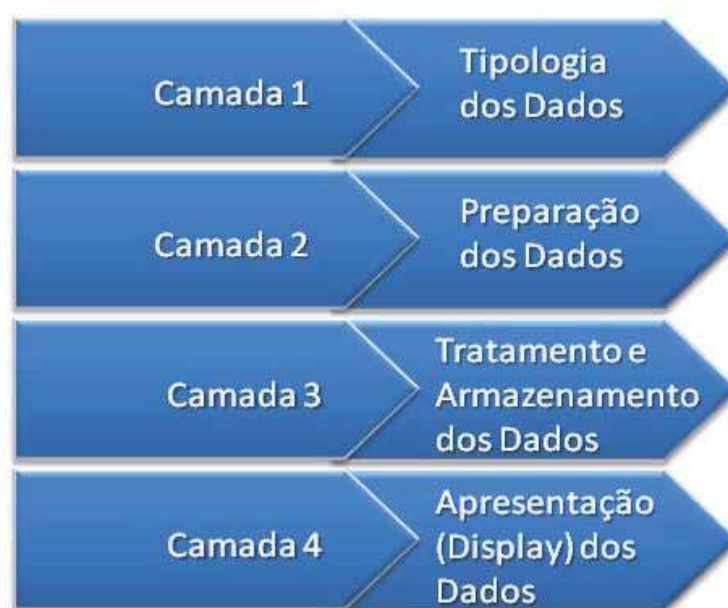
Quadro 13: Recomendações funcionais para bibliotecas digitais e *Web Semântica*.

FUNÇÃO	BIBLIOTECAS DIGITAIS	WEB SEMÂNTICA
Elementos de dados	Definidos por AACR2, ISBD e MARC 21.	Definidos por RDA, RDF, RDFS, OWL e SKOS.
Valor dos dados	Definidos por AACR2, RDA, ISBD, LC/NACO, LCSH, CDD e CDU.	Definidos por esquemas de descrição usando RDF/RDFS/OWL/SKOS
Codificação ou rotulagem de elementos de dados para manipulação por máquinas	MARC 21, FRBR e RDF.	Definidos por RDF/RDFS/XML
Estrutura de dados	Definidos por FRBR, FRAD, FRSAD, RDF, MARC 21, OWL e ontologias.	Definidos por RDF, RDFS, OWL, SKOS e DCMI <i>Abstract Model</i> .
Codificação de dados sobre os relacionamentos das entidades	MARC 21, FRBR, FRAD, RDF e ontologias.	Realizado por RDF, RDFS, OWL e SKOS na forma de links URI.
Display	MARC 21	Camadas de aplicação

Fonte: Elaborado pelo autor.

Como um ponto de partida para garantir a interoperabilidade, esta investigação, a partir das observações e reflexões no campo da Catalogação Descritiva, destaca alguns requisitos e diretrizes funcionais que podem ser utilizados no estabelecimento da interoperabilidade em ambientes informacionais digitais de uma forma mais efetiva. Para tanto, pretende-se compreender a estrutura intangível, numa proposta apresentada em camadas sobrepostas, conforme a figura 27, uma vez que elas devem trabalhar em sinergia para a consistência e funcionamento pleno do ambiente digital.

FIGURA 27: Diretrizes funcionais para a interoperabilidade.



Fonte: Elaborado pelo autor.

- Camada 1 - Tipologia dos dados: nessa fase inicial o projetista (catalogador) define quais os dados serão trabalhados para alimentação e modelagem do ambiente informacional, a partir do recurso bibliográfico a ser catalogado. Exemplo: dados textuais, imagéticos, áudio etc. Vale destacar que nessa tese nos teremos apenas nos dados textuais, explicitados nos códigos de catalogação (AACR2) e nos padrões de metadados (MARC 21).
- Camada 2 - Preparação dos dados: uma vez definidos os dados bibliográficos a serem utilizados no sistema, a preparação dos dados consiste na adoção de ferramentas para a

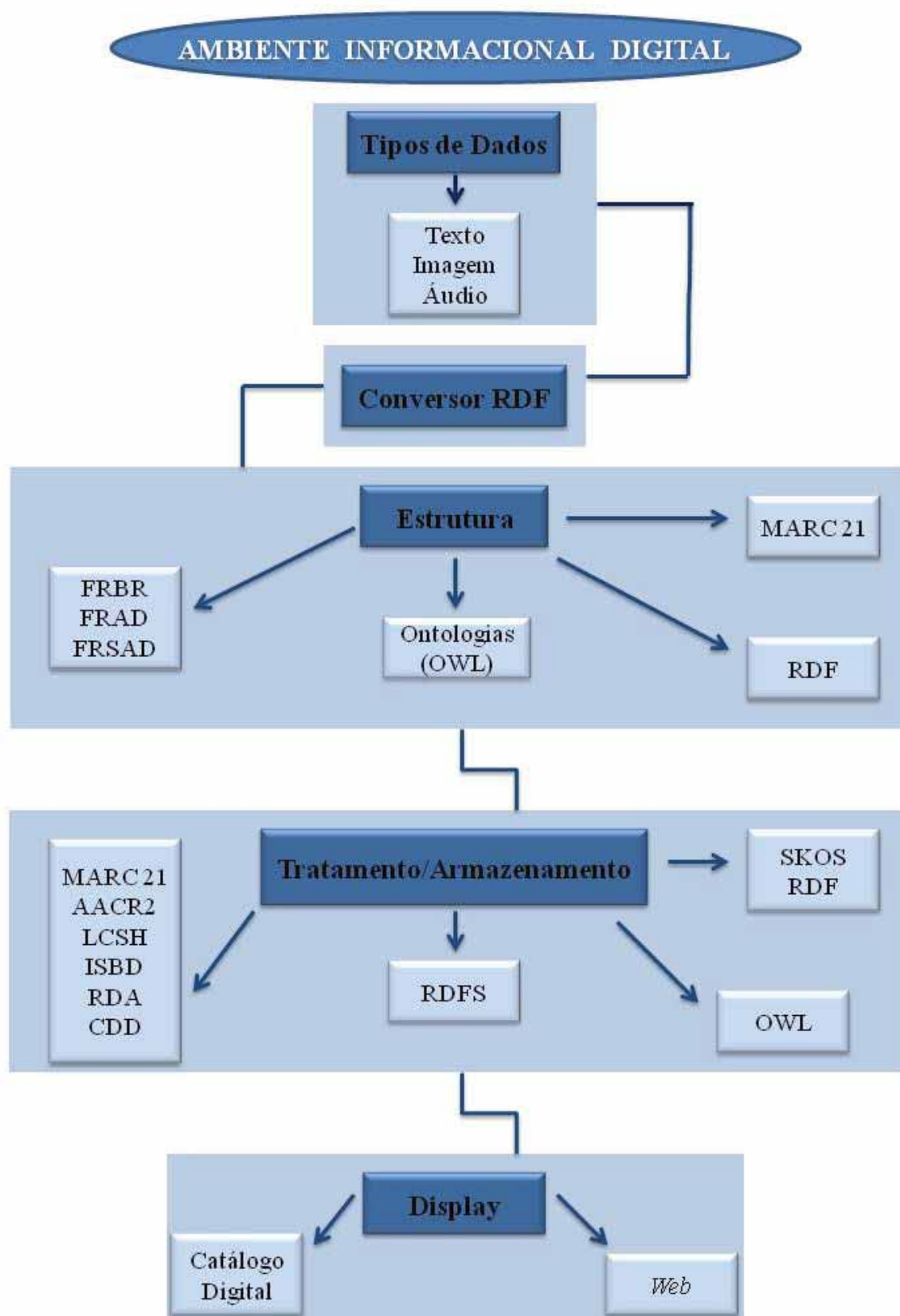
conversão⁵³ de dados em RDF. Os dados que foram extraídos de outras fontes, ou seja, dados não RDF, o W3C tem recomendado alguns conversores que auxiliam a transformação, como RDFizer. A adoção do RDFizer se justifica uma vez que não se tem oficialmente representações de dados MARC em RDF, nesse sentido, tal conversor transforma dados do padrão de metadados MODS em RDF, iniciativa essa promovida pela *Library of Congress* (MODS, 2011).

- Camada 3 – Tratamento e Armazenamento dos dados: Após a conversão dos dados em RDF, a próxima camada consiste na adoção efetiva pelo catalogador das regras e/ou dos esquemas de descrição bibliográfica (AACR2 e RDA), ou seja, a catalogação dos recursos bibliográficos, na confecção padronizada de metadados; a definição dos padrões de metadados (MARC 21), da arquitetura de metadados RDF para a estruturação dos dados e RDF *Schema* para sua validação. Vale dizer que nesse momento o catalogador deverá também adotar os Requisitos Funcionais para Registros Bibliográficos (FRBR), em consonância com os Requisitos Funcionais para Dados de Autoridade (FRAD) e os Requisitos Funcionais para Dados de Assunto (FRSAD). As ontologias aparecem nesse contexto para definir os conceitos dos elementos de um registro bibliográfico, por meio das regras e esquemas de descrição para a confecção metodológica dos metadados e dos padrões de metadados.
- Camada 4: Apresentação (*Display*) dos dados: A fase final consiste em tornar disponíveis os dados (*output*) e apresentá-los aos usuários do ambiente informacional. Os dados poderão aparecer da maneira como foram construídos e armazenados (*input*) nas camadas 1 e 3, para a camada tangível de recuperação e também visualizado na *Web*.

Acredita-se que esses requisitos e recomendações podem propiciar uma modelagem dos catálogos melhor estruturados, para posterior recuperação, uso e re (uso) das informações, garantindo a interoperabilidade e potencializando os relacionamentos bibliográficos semânticos, iniciativa essa que vai de encontro aos ideais vislumbrados pela *Web Semântica*. Para tanto, pode ser visualizado com maiores detalhes na figura 28:

⁵³ “Um conversor para RDF é uma ferramenta que converte aplicações de dados de um formato específico em RDF para uso com ferramentas de RDF e integração com outros dados. Conversores podem ser parte de um esforço de migração, ou parte de um sistema em execução que fornece uma visão *Web Semântica* de uma determinada aplicação”. (W3C, 2012, tradução nossa).

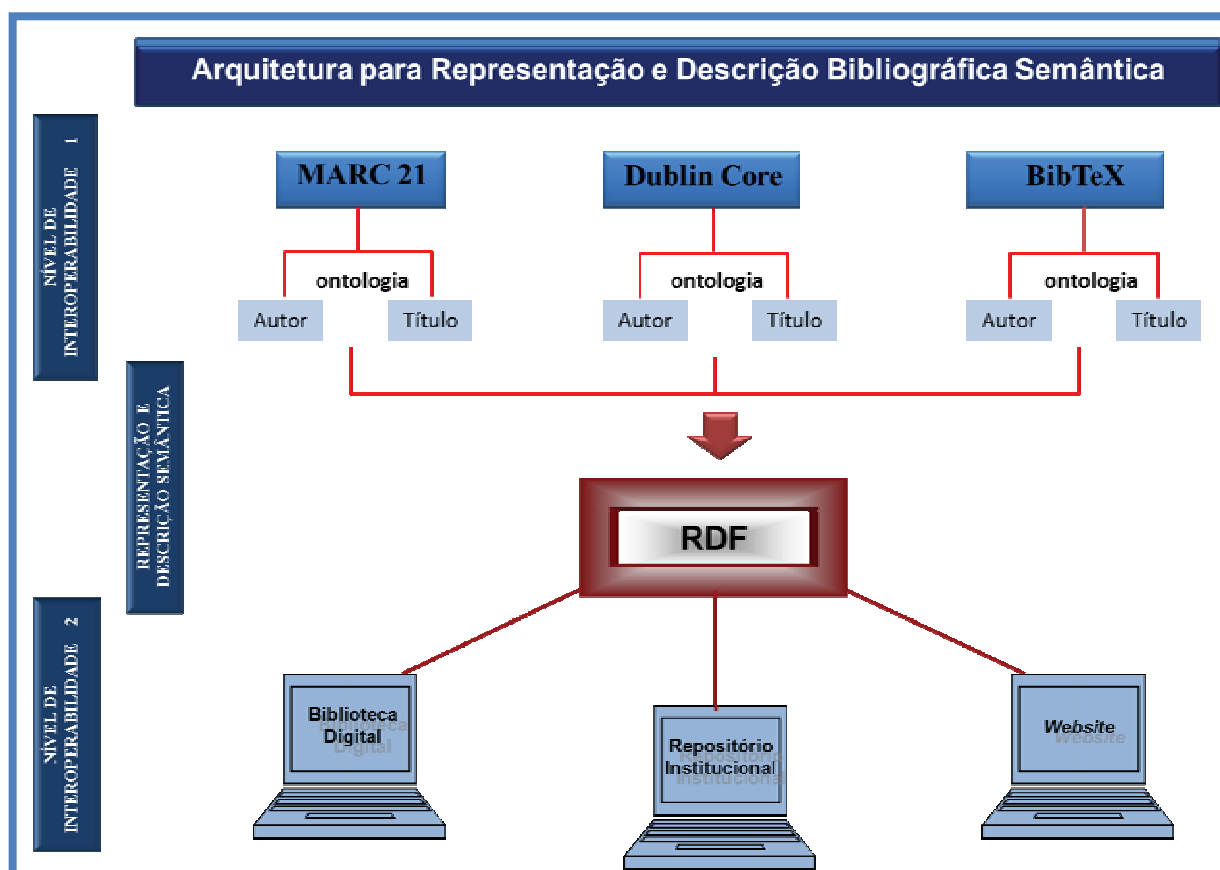
FIGURA 28: Proposta de modelagem para interoperabilidade no domínio bibliográfico.



Fonte: Elaborado pelo autor.

Com base na literatura científica apresentada sobre a representação, a descrição, a modelagem e a interoperabilidade de recursos bibliográficos e pautado nas principais definições, arquiteturas e tecnologias, a figura 29 apresenta uma proposta de um modelo/esquema para uma possível arquitetura para a representação e a descrição bibliográfica semântica de um recurso informacional, contempladas num único ambiente, com vistas a possibilitar a interoperabilidade semântica entre padrões de metadados e ambientes informacionais heterogêneos.

FIGURA 29: Arquitetura para a Representação e a Descrição Bibliográfica semântica e níveis de interoperabilidade.



Fonte: Elaborado pelo autor.

A arquitetura para a Representação e a Descrição Bibliográfica Semântica vislumbra a possibilidade da construção de representação e de descrição bibliográfica semântica e a interoperabilidade semântica entre, não somente padrões de metadados, mas também entre ambientes e sistemas informacionais digitais heterogêneos.

Na parte superior da arquitetura, têm-se alguns padrões de metadados (MARC21, Dublin Core e BibTeX), descrevendo alguns metadados em comum, tais como, autor e título (considerados elementos de calção de um recurso bibliográfico). Na fase da definição dos metadados precisa-se considerar para estruturar os elementos as ontologias, ou seja, o delineamento conceitual de cada atributo e os modelos conceituais de dados, tais como os FRBR e FRAD.

No padrão de metadados e/ou formato de intercâmbio MARC 21, por exemplo, as ontologias são determinadas através das regras e esquemas de descrição bibliográfica (AACR2) e também, do conceito fornecido pelo próprio MARC 21.

Quadro 14: Explicitação de ontologia para os elementos autor e título em AACR2.

Ontologia para o esquema descrição bibliográfica em AACR2	
Área	Ontologia
Autor	A pessoa a quem cabe à responsabilidade principal pela criação do conteúdo intelectual ou artístico de uma obra.
Título	Palavra, frase, caractere, ou grupo de caracteres, que normalmente aparecem num item, dando nome a estes ou à obra nele contida.

Fonte: Elaborado pelo autor a partir do AACR2 (Glossário D).

A definição dos elementos de metadados pela ontologia de descrição bibliográfica é considerada uma das principais partes da interoperabilidade por poder fornecer informações significativas para que sejam compreensíveis por agentes não humanos (computador/máquina).

Quadro 15: Explicitação de ontologia para elementos autor e título em MARC 21.

Ontologia para o formato MARC 21	
Etiqueta	Ontologia
100 (Autor)	Nome pessoal usado como uma entrada principal em um registro bibliográfico. A entrada principal é atribuída de acordo com as várias regras de catalogação, usualmente para o pessoal responsável pela criação intelectual de uma obra.
245 (Título)	Título e indicação de responsabilidade são áreas da descrição bibliográfica de uma obra. A indicação do título consiste da propriedade do título e também pode conter a designação geral do material (Meio), complemento do título, outras informações sobre o título, a complementação do título da página de rosto, e a indicação de responsabilidade. A propriedade do título inclui o título abreviado (<i>short title</i>) e o título alternativo (<i>alternative title</i>), a indicação numérica de uma parte/seção e o nome de uma parte/seção.

Fonte: Elaborado pelo autor a partir da *Library of Congress* (2011).

Percebem-se conceitos semânticos na ontologia de descrição bibliográfica AACR2 e MARC 21, pautados na lógica de descrição estabelecidos no processo da Catalogação Descritiva, uma vez que a base de conhecimento do formato MARC 21 é o AACR2. No RDA, por exemplo, não é explicitado o conceito Autor, o mesmo é chamado por *Creator* (Criador), linguagem que está em consonância com o padrão de metadados Dublin Core, desenhado para a localização de recursos na *Web*.

Vale destacar nesse momento que a arquitetura supracitada apresenta ontologias para apenas dois elementos da descrição bibliográfica na estrutura MARC 21 (autor e título), uma vez que é necessário o desenvolvimento e a criação de ontologias para determinar a semântica de todos os metadados do padrão MARC 21, pois quanto mais complexa e completa for a estruturação, haverá uma compreensão mais intuitiva e efetiva pela máquina promovendo a interoperabilidade.

Ainda na fase de confecção dos metadados, a atribuição das ontologias de descrição bibliográfica não é condição suficiente para determinar a interoperabilidade. O catalogador

além de conhecer a estrutura MARC 21 deverá também dominar e refletir o uso das regras de catalogação (AACR2) para saber aplicá-las no preenchimento do conteúdo dos metadados (valores dos elementos).

Para a descrição bibliográfica de um recurso bibliográfico (livro, por exemplo) na estrutura MARC 21, o catalogador deverá seguir as regras do capítulo 2 do código de catalogação (AACR2) que norteiam a construção padronizada e unívoca de como deverão ser representados e apresentados os dados aos usuários. Os metadados título e autor devem seguir as seguintes instruções:

Quadro 16: Regras de conteúdo definidas por AACR2 para elementos autor e título em MARC 21.

Regras de construção padronizada dos valores dos metadados em MARC 21, segundo o AACR2.	
Regra	Definição
2.1B Título	Transcreva o título principal de acordo com as instruções de 1.1 B.
1.1 B1.	Transcreva o título principal exatamente no que respeita à redação, ordem e grafia, mas não necessariamente quanto à pontuação e ao uso de maiúsculas. Use os acentos e outros sinais diacríticos que estiverem na fonte principal de informação (veja também 1.0 G). Use maiúsculas de acordo com o apêndice A.
2.1 F.	Indicação responsabilidade
2.1 F1.	Transcreva indicações de responsabilidade relativas às pessoas ou entidades de acordo com as instruções de 1.1 F.
1.1 F1.	Transcreva indicações de responsabilidade que figurem com destaque no item na forma em que eles aparecem. Se uma indicação de responsabilidade for extraída de uma fonte que não seja a fonte principal de informação, coloque-a entre colchetes.

Fonte: Elaborado pelo autor a partir do AACR2 (2005).

É através do uso adequado das regras e dos esquemas que as representações terão maior consistência na camada intangível do sistema propiciando formas de apresentação dos recursos bibliográficos de forma mais efetiva.

A meta no nível de interoperabilidade 1, de acordo com a figura 29, é a possibilidade de integração e comunicação entre padrões de metadados distintos, para tanto só será possível a partir da determinação de ontologias de descrição bibliográfica explicitadas nas regras e códigos de catalogação e em formatos de metadados potencializando os relacionamentos bibliográficos semânticos e estabelecendo a interoperabilidade.

A camada RDF terá o objetivo de fazer o armazenamento e a representação dos dados, localizada na parte central da arquitetura, com as descrições bibliográficas semânticas. Independente da estrutura original do padrão permitirá que os dados em MARC 21, Dublin Core, BibTeX ou qualquer outro padrão, seja convertido em RDF e podendo ser reaproveitados na construção *Web Semântica*.

Já na parte inferior da arquitetura é possível verificar a presença de ambientes informacionais heterogêneos, que podem ser interpretados tais como, Biblioteca Digital, Repositório Institucional, *Website* etc., definidos no modelo, sendo que a representação e a descrição dos recursos informacionais, podem estar ou não contemplados num mesmo padrão. O nível de interoperabilidade 2 objetiva a possibilidade de interoperabilidade semântica entre múltiplos ambientes e sistemas informacionais digitais e o armazenamento das descrições bibliográficas semânticas numa única plataforma/interface.

Ressalta-se, que quanto mais completa e complexa for uma descrição, melhores possibilidades nos níveis de interoperabilidade semântica haverá, com a possibilidade de uso, de preservação e de (re)uso das informações de forma mais efetiva pelo usuário final, multidimensionando as formas de acesso aos recursos informacionais.

Portanto, acredita-se tal modelo/esquema de arquitetura proposto nessa tese, pode possibilitar o compartilhamento entre padrões de metadados e ambientes e sistemas informacionais distintos, trabalhando numa filosofia de colaboração entre os recursos informacionais disponíveis e as tecnologias que estão abarcadas na sua construção, no

estabelecimento da interoperabilidade, na otimização dos relacionamentos bibliográficos e ampliados para a construção padronizada de recursos na *Web*.

Assim, o capítulo seguinte apresenta as considerações finais, bem como as impressões, as observações e as reflexões propostas pela pesquisa, e os novos horizontes acerca dos ambientes informacionais digitais e das questões sobre interoperabilidade de recursos bibliográficos.



CAPÍTULO 6



CONSIDERAÇÕES FINAIS

6 CONSIDERAÇÕES FINAIS

É preciso que eu, incessantemente, mergulhe na água da dúvida. *Wittgenstein*

A pesquisa denominada *Elementos de interoperabilidade na Catalogação Descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais*, pretende *a priori*, selecionar na literatura científica no campo da Biblioteconomia e da Ciência da Informação, fundamentação teórica e metodológica para responder aos requisitos que garantam a modelagem de ambientes informacionais digitais melhor estruturados para a recuperação da informação, a partir dos instrumentos de representação e de descrição das metodologias cristalizadas da Catalogação Descritiva e dos padrões de metadados, no domínio bibliográfico.

A evolução das bibliotecas tem merecido grande destaque, ao longo destes anos, no que se refere ao desenvolvimento e uso de tecnologias, principalmente as da informação e comunicação (TIC), pois potencializaram seus serviços com o oferecimento de novos recursos de acesso, com a utilização de formatos de intercâmbio de dados na Catalogação Descritiva, no uso de ontologias, na orientação para a modelagem de catálogos e no processo de busca e recuperação da informação.

A proposta MarcOnt de acordo com essa investigação científica, com base no *corpus* teórico tem em seus princípios a interoperabilidade de registros bibliográficos oriundos de diferentes ambientes informacionais digitais e com diferentes estruturas de representação (metadados ou padrões de metadados) e apresentação de recursos.

Enquanto isso, o padrão de metadados ou formato de metadados MARC 21 tem em seu princípio o intercâmbio de informações bibliográficas e catalográficas, ou seja, um padrão de domínio específico da Biblioteconomia, elaborado para conceder um modelo estruturado e rico de descrições bibliográficas, nos ambientes informacionais digitais.

Após o levantamento exaustivo da literatura científica, bem como a comunicação com o idealizador do MarcOnt, Sebastian Kruk, o pesquisador constatou que os estudos sobre a ontologia foram direcionados para outras temáticas de interesse do autor (devido sua formação na área de Ciência da Computação), tais como usabilidade em bibliotecas digitais, aspectos tecnológicos da *Web Semântica* etc.

O MarcOnt foi o objeto eleito para estudo, pois propagava ser um instrumento tecnológico que permitiria a interoperabilidade entre padrões de metadados heterogêneos. Assim, foi comprovada a necessidade de se investigar a necessidade de elementos de interoperabilidade no processo de catalogação.

Destaca-se que a iniciativa MarcOnt da maneira como foi pensada inicialmente por seu idealizador é muito relevante para o domínio bibliográfico, uma vez que pretende-se permitir a interoperabilidade entre padrões de metadados heterogêneos. Assim, a ontologia MarcOnt pode ser considerada como uma iniciativa que poderia ser aproveitada para a Catalogação Descritiva e para a interoperabilidade entre ambientes informacionais digitais.

Com base no levantamento bibliográfico realizado, verifica-se, no panorama atual da Catalogação Descritiva, que os modelos conceituais propostos pelos Requisitos Funcionais para Registros Bibliográficos (FRBR) e as bases ontológicas dos esquemas de metadados podem propiciar a clareza nos relacionamentos entre registros bibliográficos.

Vale dizer que o modelo FRBR facilita o desenho de um modelo conceitual, em consonância com as ontologias e os metadados, não somente por sua riqueza conceitual e estrutural, mas porque se constitui num marco de referência para a compreensão dos relacionamentos bibliográficos e na modelagem dos catálogos digitais.

As ontologias e os metadados apresentam-se hoje como ricos instrumentos para a representação e a descrição de recursos bibliográficos nos ambientes informacionais digitais atualmente, proporcionando, se utilizados em conjunto, a interoperabilidade semântica entre diferentes sistemas e plataformas.

A palavra de ordem num cenário marcado pela heterogeneidade de recursos informacionais, armazenados nas mais diversas mídias, é interoperabilidade, ou seja, a criação

de uma linguagem comum que permita aos sistemas e aos ambientes informacionais digitais intercambiar recursos bibliográficos.

No entanto, há uma grande quantidade de atividades em andamento para criar relacionamentos semânticos na rede, compatíveis com a *Web Semântica*. Isto requer elementos expressos de vários padrões de metadados utilizando RDF. Mais especificamente, elementos relacionados à estrutura de metadados, tais como as etiquetas, campos, subcampos e atributos, necessitam ser expressos como classes e propriedades em RDF *Schema*; enquanto os elementos relacionados ao conteúdo dos metadados, como os códigos e vocabulários controlados necessitam ser expressos em SKOS e os relacionamentos semânticos na linguagem de ontologia OWL.

O modelo RDF subjacente às tecnologias da *Web Semântica* é frequentemente descrito como o futuro de metadados estruturados. A sua adoção em bibliotecas tem sido lenta, justificado pelo fato de que as diferenças fundamentais estão na abordagem de modelagem que leva RDF, representando uma arquitetura "*bottom up*" onde as descrições são distribuídas e pode ser feita sob qualquer característica considerada necessária, enquanto a abordagem "cêntrica" de registros bibliográficos das bibliotecas tende a ser mais "*top down*" confiando nas funções pré-definidas determinadas por elas. (LEE; JACOB, 2011; YEE, 2009; RILEY, 2010; COYLE, 2012).

MARC 21 fornece o protocolo pelo qual os computadores realizam o intercâmbio, o uso e a interpretação da informação bibliográfica, ligando elementos de dados para formar a base da maioria dos catálogos de bibliotecas utilizada hoje. A *Web Semântica* em muitos aspectos será baseada na utilização desse tipo de ligação de dados, mas talvez em um catálogo ou banco de dados muito maior e globalizado.

Se o objetivo final de um projeto de biblioteca digital é organizar uma ampla gama de recursos informacionais, independentemente de sua estrutura original, para atingir a interoperabilidade é necessário a adoção de arquiteturas de metadados (RDF), dos requisitos funcionais para registros bibliográficos e dos metadados para a modelagem dos ambientes e o uso de ontologia de descrição bibliográfica, que garantirá a padronização da representação para posterior recuperação, uso e (re) uso de informações.

Modelos ontológicos de relacionamentos bibliográficos têm vantagens para a estruturação de novos catálogos. Uma idéia com muitas variações é estender campos do formato MARC 21 para ligações explícitas para outras obras. MARC 21 tem uma organização de arquivos de dados muito complexa. É possível converter a estrutura idiossincrática de registros MARC, ou seja, os códigos de campos, os subcampos e os valores condicionados aos indicadores para mais formatos padrão. A necessidade de fazê-lo, portanto, impede a utilização do avanço na corrente principal de tecnologias de banco de dados. (LEE; JACOB, 2011).

A redundância de dados reduz a eficiência do armazenamento e da atualização de um ambiente informacional. A ameaça mais séria para as bibliotecas digitais, no entanto, é a perspectiva de sobrecarregar os usuários com numerosas versões de obras muito semelhantes, no processo de busca e recuperação da informação, como acontece no ambiente *Web* atualmente.

O uso dos modelos conceituais de dados, das arquiteturas de metadados e das ontologias redesenham os novos ambientes informacionais digitais, definindo conceitualmente os elementos da descrição bibliográfica a serem representados pelo catalogador, proporcionando interfaces de buscas mais compreensíveis aos usuários e no estabelecimento efetivo da interoperabilidade.

A adoção do padrão de metadados não é condição suficiente para atender aos requisitos de interoperabilidade. Faz-se necessário a utilização correta das regras e/ou esquemas de descrição, pautados numa lógica descritiva, para a definição e construção padronizada dos metadados gerando representações dos recursos bibliográficos consistentes e unívocas, alcançando dessa forma, as potencialidades da interoperabilidade em ambientes informacionais digitais.

Acontece uma releitura dos relacionamentos bibliográficos, uma vez que os mesmos sempre existiram, porém, não explicitados. Os modelos conceituais e as ontologias para descrição bibliográfica na modelagem dos catálogos digitais fornecem a semântica explicitada através dos códigos, regras e esquemas de descrição, pautados numa lógica descritiva, na definição e construção padronizada de metadados e garante a interoperabilidade em ambientes informacionais digitais.

Num panorama geral busca-se apontar como os instrumentos de descrição bibliográfica, juntamente com uma ontologia no domínio bibliográfico, explicitada nos códigos e padrões de metadados podem propiciar a clareza nos relacionamentos entre esquemas de metadados heterogêneos influenciando decisivamente na constituição e na modelagem de novos ambientes informacionais digitais, nos contextos da *Web Semântica*, *Web 2.0* e *Web 3.0*, e na promoção da interoperabilidade semântica, pois acredita-se que essa será a tendência entre os sistemas informacionais digitais.

A utilização de ontologias, de metadados e dos fundamentos teóricos e metodológicos da Catalogação Descritiva uma vez articuladas, podem redesenhar novas opções aos ambientes informacionais, seja na modelagem do banco de dados (catálogos), ou ainda, na forma de representação dos recursos informacionais, garantindo a possibilidade de interoperabilidade semântica e no auxílio às formas de apresentação dos recursos, proporcionando aos usuários uma multiplicidade nas formas de busca, de acesso e de recuperação de informações pertinentes, significativas e relevantes, bem como o uso, a preservação e o (re) uso numa única interface.

Quanto aos profissionais da Ciência da Informação, em especial os bibliotecários catalogadores, não podem prescindir de compreender a necessidade de construção de registros bibliográficos, que estejam pautados numa base teórico-epistemológica e metodológica que os sustentem e obedeça a registros pré-determinados e estabelecidos por regras e/ou esquemas de descrição bibliográfica internacionais, visando o controle bibliográfico universal.

A estrutura estratégica para promoção da interoperabilidade em ambientes informacionais digitais se dá na união da camada intangível (representação e descrição dos dados) de instanciamento e persistência dos dados bibliográficos, com a camada visível de apresentação para os usuários. O catalogador projeta por antecipação o ambiente informacional, utilizando ontologias para a descrição bibliográfica explicitadas nas regras e esquemas para a representação e descrição dos recursos bibliográficos, definindo conceitualmente os elementos, juntamente com os modelos conceituais e os padrões de metadados, garantindo a unicidade dos recursos, apresenta-os conforme as solicitações das estratégias de busca e de recuperação ao usuário.

Nesse sentido, a garantia de interoperabilidade em ambientes informacionais digitais no domínio da Catalogação Descritiva e que potencializaria a modelagem, a estruturação e a representação de recursos bibliográficos podem ser definidos por:

- Modelos conceituais de dados: FRBR, FRAD e FRSAD
- Arquitetura de metadados: RDF e RDFS
- Ontologia: OWL
- Metadados
- Padrões de metadados: MARC 21
- Regras/esquemas de descrição bibliográfica: AACR2, ISBD's e RDA.

Acredita-se que a arquitetura para a representação e a descrição bibliográfica semântica e os níveis de interoperabilidade desenvolvida nessa tese, propicia a modelagem estrutural dos ambientes informacionais digitais atualmente, a partir da heterogeneidade de esquemas de metadados, tornando-se, dessa forma, uma estrutura única a ser adotada por vários catalogadores, ampliando seu escopo para a construção padronizada de recursos na *Web*. Além disso, a adoção desse modelo configura o estabelecimento efetivo da interoperabilidade, principalmente com os impactos tecnológicos na Catalogação Descritiva.

Espera-se com essa tese fornecer insumos teórico, epistemológico e metodológico para que os profissionais da Ciência da Informação revisitem a Catalogação Descritiva, não enquanto uma simples técnica, mas enquanto um processo inteligente e reflexivo que determina a modelagem dos catálogos digitais, oferecendo interfaces para a recuperação de informações mais consistentes e significativas para o usuário final.

Quanto ao “futuro” do MARC 21, sabe-se que há iniciativas internacionais de construção de conversores, ferramentas e instrumentos tecnológicos que podem potencializar a transformação dos dados bibliográficos, numa estrutura RDF, mas ainda nenhuma oficializada pelos organismos internacionais de catalogação. Essa tese não tem a pretensão de desenvolver o que poderia ser chamado de MARC RDF, mas refletir sobre um caminho inicial com orientações teóricas e metodológicas, para a determinação e a construção de ambientes informacionais melhor estruturados, com descrições bibliográficas semânticas explicitadas, propiciando a interoperabilidade em ambientes informacionais digitais.



REFERÊNCIAS



REFERÊNCIAS

AACR2. *Anglo American Cataloging Rules*. Preparado sob a direção de The Joint Steering Committee for Revision of AACR; trad. Federação Brasileira de Associações de Bibliotecários, Cientistas da Informação e Instituições (FEBAB). 2. ed., rev. 2002. São Paulo: FEBAB, 2005.

ABRAMSON, M.; MEANS, G. E. *E-Government 2001 IBM Endowment for the Bussiness of Government*. Rowman & Littlefield Publishers, Inc. 2001.

ADIDA, Ben et al. *RDFa in XHTML: syntax and processing*. W3C, Oct. 2008. Disponível em: <<http://www.w3.org/TR/2008/REC-rdfa-syntax-20081014/>>. Acesso em: 22 maio 2010.

ALBUQUERQUE, Nokolai Dimitrii; KERN, Vinícius Medina. Uma arquitetura de compartilhamento de conhecimento em bibliotecas digitais. In: SEMINÁRIO DE COMPUTAÇÃO – SEMINCO, 13., 2004. *Anais...* Blumenau: Universidade Regional de Blumenau, 2004. Disponível em: <<http://www.inf.furb.br/seminco/2004/artigos/120-vf.pdf>>. Acesso em: 18 dez. 2011.

ALMEIDA, Maurício B. *Um modelo baseado em ontologias para representação da memória organizacional*. 2006. 345 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

ALMEIDA, Maurício B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. Brasília, *Ciência da Informação*, v. 32, n. 3, p. 7-20, set./dez. 2003.

ALVARENGA, Lídia. Organização da informação nas bibliotecas digitais. In: NAVAES, M. M. L.; KURAMOTO, H. (Org.). *Organização da informação: tendências e princípios*. Brasília: Briquet de Lemos/Livros, 2006. 142 p. p. 76-98.

ALVES, R. C. V. *Web Semântica: uma análise focada no uso de metadados*. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia Ciências, Universidade Estadual Paulista, Marília, 2005.

ALVES, Rachel C. V. *Metadados como elementos do processo de catalogação*. 2010. 134f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

ARMS, William. A spectrum of interoperability. *D-Lib Magazine*, Vol. 8, no. 1, Jan. 2002.

BACA, Murtha. *Introduction to metadata 3.0*. 2. ed. Los Angeles: The Getty Research Institute, 2008. Disponível em:
<http://www.getty.edu/research/conducting_research/standards/intrometadata/pdf.html>. Acesso em: 20 maio 2010.

BERNERS-LEE, Tim. *Semantic web concepts*. 2005. 32 p. Disponível em:
<<http://www.w3.org/2005/Talks/0517-boit-tbl>>. Acesso em: 18 mar. 2010.

BERNERS-LEE, T. *Semantic web - XML2000*. [2005?]. Disponível em:
<<http://www.w3.org/2000/Talks/1206-xml2k-tbl/Overview.html>>. Acesso em: 21 maio 2007.

BLATTMANN, U.; SILVA, F. C. C. da. Colaboração e interação na web 2.0 e biblioteca 2.0. *Revista ACB: Biblioteconomia em Santa Catarina*, Florianópolis, v. 12, n. 2, p. 191-215, jul./dez., 2007.

BRITISH LIBRARY. *Exchange Formats*. [S. l.: S. n.], 2003. Disponível em:
<<http://www.bl.uk/services/bibliographic/exchange.html>>. Acesso em: 17 set. 2005.

CAMPOS, Maria L. M.; CAMPOS, Maria L. de. A.; CAMPOS, Linair M. Web semântica e a gestão de conteúdos informacionais. In: MARCONDES, C. H. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. 2. ed. Salvador: EDUFBA; Brasília: IBICT, 2006. p. 55-74.

CASA, M. E. Ambientes computacionais com múltiplas formas de representação. In: GONZÁLES, M. E. Q. et al. (Org.). *Encontro com as ciências cognitivas*. 2. ed. rev. e ampl. Marília: Faculdade de Filosofia e Ciências, 1997. p. 209-220.

CASTRO, Fabiano. F. de; SANTOS, Plácida L. V. A. C. Os metadados como instrumentos tecnológicos na padronização e potencialização dos recursos informacionais no âmbito das bibliotecas digitais na era da web semântica. João Pessoa, *Informação & Sociedade: Estudos*, v. 17, n. 2, p. 13-19, maio/ago. 2007.

CASTRO, Fabiano. F. de. *Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da ciência da informação: uma abordagem do MarcOnt initiative na era da web semântica*. 2008. 201 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília, 2008.

CASTRO, Fabiano. F. de; SANTOS, Plácida L. V. A. C. Uso das tecnologias na representação descritiva: o padrão de descrição bibliográfica semântica MarcOnt initiative nos

ambientes informacionais digitais. Brasília, *Ciência da Informação*, v. 38, n. 1, p. 74-85, jan./abr. 2009.

CASTRO, Fabiano. F. de; SANTOS, Plácida L. V. A. C. Catalogação e metadados: interlocuções nos ambientes informacionais digitais. In: ARELLANO, Filiberto Felipe Martínez (Org.). *Memoria del IV encuentro de catalogación y metadatos*. México: Centro Universitario de Investigaciones Bibliotecológicas de la Univ. Nacional Autónoma de México, 2010. v. 4, p. 301-318.

CATARINO, Maria E.; BAPTISTA, Ana A. Web semântica e a qualidade no intercâmbio da informação. In: TOMAÉL, Maria I. (Org.). *Fontes de informação na internet*. Londrina: EDUEL, 2008. p. 31-51. 176 p.

CERVO, A. L.; BERVIAN, P. A. *Metodologia científica*. 5. ed. São Paulo: Prentice Hall, 2003. 242 p.

CHATEAUBRIAND, Oswaldo. A filosofia, a linguagem e o mundo. In: BRITO, Adriano Naves de; VALE, Oto Araújo. *Filosofia, linguística, informática: aspectos da linguagem*. Goiânia: Ed. UFG, 1998.

CORCHO, Oscar; FERNÁNDEZ-LOPEZ, Mariano; GÓMEZ-PÉREZ, Asunción. OntoWeb: ontology-based information exchange for knowledge management and electronic commerce. *Technical RoadMap*, v. 1.0. Madrid: Universidad Politécnica de Madrid, 2001. Disponível em: <http://babage.dia.fi.upm.es/ontoweb/wp1/OntoRoadMap/documents/D11_v1_0.pdf>. Acesso em: 23 nov. 2011.

COYLE, Karen. MARC 21 as data: a start. *Code4Lib Journal*, Vol. 14, 2011.

COYLE, Karen. *Bibliographic framework: RDF and linked data*. Berkeley, United States, Jan. 2012. Disponível em: <<http://kcoyle.blogspot.com/2012/01/bibliographic-framework-rdf-and-linked.html>>. Acesso em: 20 jan. 2012.

CRUZ, Isabel F.; XIÃO, Huiyong. The role of ontologies in data integration. *Journal of Engineering Intelligent Systems*, vol. 13, no. 4, 2005. Disponível em: <<http://www2.cs.uic.edu/~advis/publications/dataint/eis05j.pdf>>. Acesso em: 10 abr. 2010.

CUNHA, Luiz Manoel Silva. *Web semântica: estudo preliminar*. Campinas: Embrapa Informática Agropecuária, 2002. Disponível em: <<http://www.cnptia.embrapa.br/modules/tinycontent3/content/2002/doc18.pdf>>. Acesso em 03 fev. 2010.

DABROWSKI, Maciej; KRUK, Sebastian Ryszard. *MarcOnt ontology specification*. Galway, Ireland: [Digital Enterprise Research Institute](#); Poland: [Faculty of Electronics, Telecommunication and Informatics](#), Jun. 2007. Disponível em: <<http://www.marcont.org/ontology/index.html>>. Acesso em: 12 maio 2010.

DABROWSKI, Maciej; SYNAK, M.; KRUK, S. R. Bibliographic ontology. In: KRUK, S. R.; McDANIEL, Bill. (Eds.). *Semantic digital libraries*. Berlin, Heidelberg: Springer-Verlag, 2009. 235 p. p. 103-122. Disponível em: <<http://www.springerlink.com/content/j40444k3t3348l25>> Acesso em: 16 maio 2010.

DAVENPORT, T; PRUSAK, L. *Ecologia da informação: porque só a tecnologia não basta para o sucesso na era da informação*. 4. ed. São Paulo: Futura, 2001. 316 p.

DAVIS, Ian; NEWMAN, Richard. *Expression of core FRBR concepts in RDF*. 2005. Disponível em: <<http://vocab.org/frbr/core.html>>. Acesso em: 20 jan. 2012.

DELSEY, Tom. *RDA: resource description and access*. 2010. Disponível em: <<http://www.rda-jsc.org/rda.html>>. Acesso em: 25 de fev. 2012.

DEMPSEY, Lorcan; HEERY, Rachel. *A review of metadata: a survey of current resource description formats*. Mar.1997. Disponível em: <http://www.ukoln.ac.uk/metadata/desire/overview/rev_ti.htm>. Acesso em: 30 maio 2010.

FEITOSA, Ailton. *Organização da informação na web: das tags à web semântica*. Brasília: Thesaurus, 2006. 132 p.

FERNANDES, A. G. E-governo: o que já fazem estados e municípios. *Informe-se [on-line]*n. 20, out. 2000. Disponível em:<
http://www.bndes.gov.br/SiteBNDES/export/sites/default/bndes_pt/Galerias/Arquivos/conhecimento/informesf/inf_20.pdf>. Acesso em: 17 jan. 2012.

FERREIRA, M. M. (Trad. e Adap.). *MARC 21: formato condensado para dados bibliográficos*. 2.ed. Marília: Universidade Estadual Paulista, 2002. v. 1.

FOAF. *Documentation*. 2010. Disponível em: <<http://www.foaf-project.org/original-intro>>. Acesso em: 21 maio de 2010.

FURRIE, B. *Understanding MARC Bibliographic: Machine – Readable Cataloging*. 2000. Disponível em: <<http://www.loc.gov/marc/umb/>>. Acesso em: 30 jul. 2007.

GODBY, C. J.; SMITH, D.; CHILDRESS, E. Two paths to interoperable metadata. In: *Proceedings of DC-2003: Supporting Communities of Discourse and Practice. Metadata Research & Application*. Seattle, Sep./Oct. 2003. Disponível em: <http://www.siderean.com/dc2003/103_paper-22.pdf>. Acesso em: 20 jan. 2012.

GREENBERG, J.; SUTTON, S.; CAMPBELL, D. Metadata: a fundamental component of the semantic web. *Bulletin of the American Society for Information Science*, v. 29, n. 4, p.16-18, 2003.

GRUBER, Tom. Ontology. In: LIU, Ling; ÖZSU, M. Tamer (Eds.). *Encyclopedia of Database Systems*. [S.l.]: Springer-Verlag, 2009. Disponível em: <<http://tomgruber.org/writing/ontology-definition-2007.htm>>. Disponível em: 02 jan. 2012.

GRUBER, Tom et al. *Ontology framework draft statement*. 2007. Disponível em: <http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2007_FrameworksForConsideration>. Acesso em: 14 dez. 2011.

GUY, Marieke. *Interoperability focus: looking at interoperability*. Ukoln, 2005. Disponível em: <<http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>>. Acesso em: 20 mar. 2011.

HAAV, H. M.; LUBI, T. L. A survey of concept-based information retrieval tools on the web. In: EAST-EUROPEAN CONFERENCE ON ADVANCES IN DATABASES AND INFORMATION SYSTEMS, 5., Sep. 2001. *Proceedings...* Lithuania: Vilnius Gediminas Technical University, 2001. Disponível em: <<http://www.science.mii.it/ADBIS/local2/haav.pdf>>. Acesso em: 02 dez. 2011.

HEATH, Tom; BIZER, Christian. *Linked data: evolving the web into a global data space*. [Morgan & Claypool](#), Vol. 1 no. 1, 2011. 136 p. (Synthesis Lectures on the Semantic Web: Theory and Technology). Disponível em: <<http://linkeddatabook.com/editions/1.0/#linkedData>>. Acesso em: 15 nov. 2011.

HILLMANN, Diane et al. RDA vocabularies: process, outcome, use. *The Magazine of Digital Library Research*, Vol. 16, no. 1/2, Jan. /Feb. 2010.

IEEE. STANDARDS INFORMATION NETWORK. *The authoritative dictionary of IEEE standards terms*. 7 ed. New York: IEEE, 2000.

IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional requirements for bibliographic records: final report*. UBCIM Publications - New Series, vol. 19. München: K. G. Saur, 1998. 136 p. Disponível em: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>. Acesso em: 18 dez. 2011.

JACOB, E. K. Ontologies and the semantic web. *Bulletin of the American Society for Information Science and Technology*, v. 29, n. 4, p. 19-22, Apr./May, 2003. Disponível em: <<http://www.asis.org/Bulletin/Apr-03/jacob.html>>. Acesso em: 27 abr. 2010.

KAKALY, Constantia et al. *Integration dublin core metadata for cultural heritage collections using ontologies*. Singapore, Aug. 2007. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS "Application Profiles: Theory and Practice", pp. 128-139, 2007. Disponível em: <<http://dcpapers.dublincore.org/ojs/pubs/article/view/871/867>>. Acesso em: 27 abr. 2010.

KIM, Hak L. et al. The state of the art in tag ontologies: a semantic web model for tagging and folksonomies. Berlin, Sep. 2008. In: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, *Proceedings...* Berlin, 2008, pp. 128-137. Disponível em: <<http://dcpapers.dublincore.org/ojs/pubs/article/viewFile/925/921>>. Acesso em: 20 maio 2010.

KRUK, Sebastian R.; McDANIEL, Bill. (Eds.). *Semantic digital libraries*. Berlin, Heidelberg: Springer-Verlag, 2009. 235 p. Disponível em: <<http://www.springerlink.com/content/p77265/?p=b55d5c38d27d45b7814f706b9d04698c&pi=0>>. Acesso em: 21 maio 2010.

KRUK, Sebastian R. *Semantic digital libraries: improving usability of information discovery with semantic and social services*. Heidelberg: Kindle, 2010. Disponível em: <<http://www.lulu.com/product/item/semantic-digital-libraries---improving-usability-of-information/6343371>>. Acesso em: 20 maio 2010.

KRUK, S. R.; SYNAK, M.; ZIMMERMANN, K. *MarcOnt initiative: mediation services for digital libraries*. 2005a. Disponível em: <<http://www.marcont.org/marcont/pdf/marcontecd12005.pdf>>. Acesso em: 16 jan. 2006.

KRUK, S. R.; SYNAK, M.; ZIMMERMANN, K. *MarcOnt initiative: integration ontology for bibliographic description formats*. Madrid, 2005b. Disponível em: <<http://dc2005.uc3m.es/program/presentations/Thursday%2015.%2015.30%20h%20-%20s.kruk.pdf>>. Acesso em: 16 jan. 2006.

KRUK, S. R.; SYNAK, M.; ZIMMERMANN, K. *MarcOnt initiative: bibliographic description related tools utilizing semantic web technologies*. Ireland, Galway: Digital Enterprise Research Institute, Jun., 2005c. Disponível em: <http://www.marcont.org/index.php?option=com_content&task=view&id=15&Itemid=28>. Acesso em: 12 maio 2010.

LASSILA, O.; SWICK, R. R. *Resource description framework (RDF) model and syntax specification*. [S. l.: S. n.], 1999. Disponível em: <<http://www.w3.org/TR/1999/REC-rdfsyntax-19990222/>>. Acesso em: 3 jul. 2007.

LEE, Seungmin; JACOB, Elin K. An integrated approach to metadata interoperability: construction of a conceptual structure between MARC and FRBR. *Library Resources & Technical Services*, Vol. 55, no. 1, p.17-32, Jan. 2011.

LIBRARY OF CONGRESS. *MARC 21 concise formats*. 2006. Disponível em: <<http://www.loc.gov/marc/concise/concise.html>>. Acesso em: 30 jul. 2007.

LIBRARY OF CONGRESS. *MARC 21 format for bibliographic data*. Library of Congress Network Development and MARC Standards Office, No. 13, Sep. 2011. Disponível em: <www.loc.gov/marc/bibliographic/ecbdhome.html>. Acesso em: 12 jan. 2012.

LIMA-MARQUES, Mamede. *Ontologias: da filosofia à representação do conhecimento*. Brasília: Thesaurus, 2006. 69 p. (Série Ciência da Informação e da Comunicação; v. 1).

LYTRAS, M. D; SICILIA, M. A. Where is the value of metadata? *International Journal of Metadata, Semantics and Ontologies*, Vol. 2, no. 4, p. 235-241, 2007.

MARCONDES, Carlos H. et al. Ontologias como novas bases de conhecimento científico. *Perspectivas em Ciências da Informação*, v. 13, n. 13, p. 20-39, set./dez. 2008.

MARINO, M. T. *Integração de informações em ambientes científicos na web: uma abordagem baseada na arquitetura RDF*. 2001. Dissertação (Mestrado em Informática) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001. Disponível em: <http://genesis.ncl.ufrj.br/dataware/Metadados/Teses/Teresa/pagina_tese.htm>. Acesso em: 22 jan. 2011.

MATTELART, A. *História da sociedade da informação*. São Paulo: Loyola, 2002.

MÉNDEZ RODRÍGUEZ, E. *Metadados y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002. 429 p.

MILES, Alistair; BECHHOFFER, Sean. *SKOS: simple organization system namespace document – html variant*. W3C: Aug. 2009. Disponível em: <<http://www.w3.org/TR/skos-reference/skos.html#hasTopConcept>>. Acesso em: 31 maio de 2010.

- MILLER, P. *Metadata for the masses*. [S. l.: S. n.], 1996. Disponível em: <<http://www.ariadne.ac.uk/issue5/metadata-masses/>>. Acesso em: 16 abr. 2004.
- MILSTEAD, J.; FELDMAN, S. *Metadata: cataloging by any other name*. *Online*, [S. l.], January 1999. Disponível em: <<http://www.online.com/online/ol1999/milstead1.html>>. Acesso em: 22 mar. 2007.
- MODS. *Metadata Object Description Schema*. Library of Congress, Mar. 2011. Disponível em: <<http://www.loc.gov/standards/mads/rdf/>>. Acesso em: 21 maio 2011.
- NILSSON, Mikael. *From interoperability to harmonization in metadata standardization: designing an evolvable framework for metadata harmonization*. 2010. 125 f. Thesis (Doctoral in Computer Science and Communication) - KTH School of Computer Science and Communication, Stockholm, 2010.
- NISO. *Understanding metadata*. Bethesda, USA: National Information Standards Organization, 2004. 20 p. Disponível em: <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>>. Acesso em: 10 abr. 2010.
- NOVÁČEK, Vít. et al. *Extending community ontology using automatically generated suggestions*. Galway: American Association for Artificial Intelligence, 2007. Disponível em: <<http://www.aaai.org/Papers/FLAIRS/2007/Flairs07-060.pdf>>. Acesso em: 12 abr. 2010.
- OWL. Web Ontology Language. *Use cases and requirements*. W3C recommendation, 10 Feb. 2004. Disponível em: <<http://www.w3.org/TR/webont-req/#onto-def>>. Acesso em: 27 abr. 2010.
- PETERSON. D. (Org.) *Forms of representation: an interdisciplinary theme for cognitive science*. Wiltshire: Cromwell Press, 1996. 208 p.
- PICKLER, Maria. E. V. Web semântica: ontologias como ferramentas de representação do conhecimento. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 12, n. 1, p. 65-83, jan./abr. 2007.
- RAMALHO, Rogério, Ap. S. *Desenvolvimento e utilização de ontologias em bibliotecas digitais: uma proposta de aplicação*. 2010. 145 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

RDA. *Glossary*. American Library Association, Canadian Library Association, and The Chartered Institute of Library and Information Professionals, Oct. 2008.

RILEY, Jenn. *RDF for librarians*. DLP Brown Bag Series, p. 1-38, Sep. 2010. Disponível em: < <http://www.dlib.indiana.edu/education/brownbags/fall2010/rdf/rdf.pdf>>. Acesso em: 15 nov. 2011.

RILEY, Jenn. Enhancing interoperability of FRBR-based metadata. In: *Proc. Int'l Conf. on Dublin Core and Metadata Applications*. 2010a. p. 31-43. Disponível em: <<http://dcpapers.dublincore.org/index.php/pubs/article/view/1037>>. Acesso em: 12 dez. 2011.

SANTOS, Plácida, L. V. A. C.; ALVES, Rachel C. V. Metadados e web semântica para a estruturação da web 2.0 e web 3.0. *Datagramazero – Revista de Ciência da Informação*, v. 10, n. 6, dez. 2009.

SANTOS, Plácida. L. V. A. C.; CORRÊA, Rosa. M. R. *Catálogo: trajetória para um código internacional*. Niterói: Intertexto, 2009. 80 p.

SANTOS, Plácida. L. V. A. C.; [VIDOTTI, Silvana. A. B. G.](#) Perspectivismo e Tecnologias de Informação e Comunicação: acréscimos à Ciência da Informação? *DataGramaZero - Revista de Ciência da Informação*, Rio de Janeiro, v. 10, n. 3, p. 1-12, jun. 2009.

SAYÃO, Luis F.; MARCONDES, Carlos H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, Campinas, v. 20, n. 2, p.133-148, maio/ago. 2008.

SHET, A.; ARPINAR, I. B.; KHASYAP, V. *Relationships at the heart of semantic web: modeling, discovering and exploiting complex semantic relationships*. Athens: University of Georgia, 2002. Disponível em: < <http://lstdis.cs.uga.edu/lib/download/SAK02-TM.pdf> >. Acesso em: 27 abr. 2010.

SIQUEIRA, M. A. *XML na ciência da informação: uma análise do MARC21*. Marília, 2003, 134 f. Dissertação (Mestrado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

SOUZA, Renato R.; ALVARENGA, Lúcia. A web semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, v. 33, n. 1, p. 132-141, jan./abr. 2004.

SOWA, John F. Ontology, metadata and semiotics. In: INTERNATIONAL CONFERENCE ON CONCEPTUAL STRUCTURES, 8., 2000, *Proceedings...* German: B. Ganter & G. W.

Mineau, 2000. Disponível em: < <http://users.bestweb.net/~sowa/peirce/ontometa.htm>>. Acesso em: 10 abr. 2010.

St. PIERRE, Margaret; LaPLANT JR., William P. *Issues in crosswalking metadata content standards*. National Information Standards Organization, Oct. 1998. Disponível em: < http://www.niso.org/publications/white_papers/crosswalk/>. Acesso em: 21 abr. 2010.

STYLES, Rob; AYERS, Danny; SHABIR, Nadeem. *Semantic Marc, MARC 21 and the semantic web*. Knight's Court: Draft, Feb. 2008. Disponível em: < <http://events.linkeddata.org/ldow2008/papers/02-styles-ayers-semantic-marc.pdf>>. Acesso em: 12 jan. 2010.

SYNAK, M. *MarcOnt Ontology: Semantic MARC21 Description for L2L & L2C Communication*. 2005. 126 f. Master's Thesis (Informatics, Distributed Computer Systems) - Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, 2005. Disponível em: <<http://library.deri.ie/servlet/showPDF?docId=http%3a%2f%2flibrary.deri.ie%2fresource%2f3a9faf28&chapter=1&view=pdf>>. Acesso em: 25 jul. 2007.

TAYLOR, Arlene. G. *The organization of information*. 2. nd. Westport: Libraries Unlimited, 2004. 417 p. (Library and Information Science Text Series).

TILLET, B. *Designation of roles in RDA*. Memorandum to the Joint Steering Committee for Development of RDA, Feb. 5, 2008. Disponível em: <http://www.rdaonline.org/constituencyreview/Phase1AppI_10_27_08.pdf>. Acesso em: 29 nov. 2011.

THOMALE, Jason. *Interpreting MARC: where's the bibliographic data?* *Code4Lib Journal*, Vol. 11, 2010.

USCHOLD, Mike; JASPER, Robert. A framework for understanding and classifying ontology applications. In: WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING METHODS, 1., Stockholm, 1999. Proceedings... Sweden, 1999. Disponível em: <<http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/11-uschold.pdf>>. Acesso em: 29 nov. 2011.

VAN HEIJST, G.; SCHREIBER, A.; WIELINGA, B. J. Using explicit ontologies in KBS development. *Journal Human-Computer Studies*, Vol. 45, p. 183-292, 1997.

VELLUCCI, S. L. Metadata and authority control. *Library Resources and Technical Services*. Vol. 44, no. 1, p. 33-43, Jan. 2000.

VRA. Visual Resources Association. *Data standards*. 2010. Disponível em: <<http://www.vraweb.org/>>. Acesso em: 20 maio de 2010.

WORLD WIDE WEB CONSORTIUM (W3C). *OWL 2 Web Ontology Language, Structural Specification and Functional-Style Syntax*, W3C Recommendation, Oct. 2009. Disponível em: <<http://www.w3.org/TR/owl-syntax/>>. Acesso em: 20 dez. 2011.

W3C. *Converter to RDF*. 2012. Disponível em: <<http://www.w3.org/wiki/ConverterToRdf>>. Acesso em: 21 jan. 2012.

YEE, Martha M. Can bibliographic data be put directly onto the semantic web? *Information Technology and Libraries*, Vol. 28, no. 2, p. 55-80, Jun. 2009.

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

ZHANG, Xiang; CHENG, Gong; QU, Yuzhong. Ontology summarization based on RDF sentence graph. In: WORLD WIDE WEB CONFERENCE, 16., 2007, *Proceedings...* Banff, Canada, 2007, pp. 707-715. Disponível em: <<http://www2007.org/papers/paper565.pdf>>. Acesso em: 20 maio 2010.