



RACHEL CRISTINA VESÚ ALVES

Metadados como elementos do processo de catalogação

Marília
2010



RACHEL CRISTINA VESÚ ALVES

Metadados como elementos do processo de catalogação

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação, como exigência 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: Informação e Tecnologia.

Orientador(a): Dr.^a Plácida Leopoldina Ventura Amorim da Costa Santos.
Pesquisa financiada pela CAPES.

Marília
2010

A474a Alves, Rachel Cristina Vesu.
Metadados como elementos do processo de catalogação /
Rachel Cristina Vesú Alves. -- Marília, R. C. V. Alves, 2010.
132 f. ; 30 cm.

Tese (Doutorado em Ciência da Informação) – Faculdade de
Filosofia e Ciências – Universidade Estadual Paulista, 2010.
Bibliografia: f. 127-132.
Orientadora: Profª Drª Plácida Leopoldina Ventura
Amorim da Costa Santos.

1. Metadados. 2. Padrões de metadados. 3. Tratamento
descritivo da informação. I. Autor. II. Título.

CDD 025.316

Rachel Cristina Vesú Alves

Metadados como elementos do processo de catalogação

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação, como exigência 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: Informação e Tecnologia.

Prof^a. Dr^a. Plácida L. V. A. da Costa Santos (Orientadora)
Faculdade de Filosofia e Ciências – UNESP

Prof^a. Dr^a. Silvana Ap. B. Gregório Vidotti
Faculdade de Filosofia e Ciências - UNESP

Prof. Dr. Ricardo Cesar Gonçalves Santana
Faculdade de Filosofia e Ciências - UNESP

Prof. Dr. Marcos Luiz Mucheroni
CBD-ECA-USP

Prof.^a Dr^a. Isa Maria Freire
Universidade Federal da Paraíba

Marília, 15 de março de 2010.

***“Eu sou aquela mulher que fez a escalada
da montanha da vida removendo pedras e plantando flores”***

(Cora Coralina)

DEDICATÓRIA

A Deus e aos anjos presentes em meu caminho
Aos meus pais Bene e Mauro
À minha irmã Roberta

Agradecimentos

A Deus, pela força e inspiração, por não me desamparar nos momentos difíceis, por permitir a conquista de mais um sonho. Agradeço aos Anjos em minha vida, visíveis e invisíveis.

Expresso a minha eterna gratidão àqueles que sempre estiveram presentes em todos os momentos dessa caminhada: meus pais Bene e Mauro e minha irmã Roberta. Obrigada pelo apoio, pelas palavras de incentivo, pela paciência, pela ajuda e força, por todo o amor e carinho, por sonharem comigo. Esta conquista só foi possível com o apoio de vocês. Obrigada! Amo vocês todos!!!

À Prof.^a Dr.^a Plácida L.V.A. da Costa Santos, pelos ensinamentos e conhecimentos compartilhados durante esses anos de pesquisa, pelo apoio, incentivo, pela confiança e compreensão, pois sem isso não teria chegado até aqui. Agradeço principalmente pelo carinho e amizade. Obrigada por tudo!

Agradeço aos professores da banca pela disposição e disponibilidade em participar da defesa desta tese, em especial à professora Silvana Ap. B. G. Vidotti e ao prof. Ricardo Santana, pelas contribuições durante a qualificação.

Aos amigos do laboratório de Tecnologias Informacionais, pela convivência e pela enriquecedora troca de conhecimento durante esses anos.

Aos docentes do curso de Pós-Graduação em Ciência da Informação, pelos valiosos ensinamentos transmitidos.

Ao pessoal da Pós-Graduação, colegas de turma e funcionários.

À CAPES, pelo financiamento desta pesquisa.

À Prof.^a Yolanda Matsuda pela revisão do texto, disponibilidade e gentileza.

Aos familiares que sempre torceram por mim: minha avó, meus tios, irmãos, primos e sobrinhos. Em especial agradeço à Renata, minha irmã do coração, ao André e às minhas lindinhas Juliana e Lívia, por me receberem de braços abertos em São Paulo. Obrigada por tudo!

Aos meus queridos amigos: Fabiana Straioto, Mara Patrícia, Elvis Rossi, Fabiano, Lourdes Mariano, Lucilene, Maria José, Aldinar, por todos os momentos compartilhados, pelo apoio, pelos momentos de desabafo, pelas alegrias, pelas risadas... Obrigada por tudo!

Agradeço aos novos amigos, em especial à D. Rosa, pelas orações e por ajudar a reencontrar a minha fé.

E por fim agradeço a todas as pessoas que direta ou indiretamente contribuíram para a realização deste trabalho.

Muito Obrigada!!!

RESUMO

O início do século XXI vem sendo marcado por mudanças relacionadas à informação e à tecnologia, principalmente no que diz respeito à produção e ao uso da informação, nas tecnologias de informática e desenvolvimento das TICs (Tecnologias de Informação e Comunicação) utilizadas para difundir essas informações. Diante disso, a representação da informação passou a ser requisito em diversas áreas do conhecimento, e para atender a essa necessidade foram criados vários padrões de metadados, que variam desde estruturas simples, passando por um tipo de padrão intermediário, até padrões de estruturas mais complexas de descrição. Com a variedade de tipos de padrões, a área de Ciência da Informação se depara com o seguinte impasse que se constitui como problema de pesquisa: como construir representações adequadas e padronizadas que garantam a unicidade e recuperação eficiente dos recursos informacionais? Esta pesquisa, que pertence à linha Informação e Tecnologia, apresenta como tema a análise das Formas de Representação dos recursos informacionais, tendo em vista as mudanças causadas pelos avanços das TICs no tratamento descritivo da informação (TDI) com o surgimento dos metadados. O foco constitui-se na análise das teorias, dos princípios, dos fundamentos e dos instrumentos que norteiam a construção de representações padronizadas na área de Ciência da Informação. A tese levantada para esta pesquisa é de que a integração estratégica entre as tecnologias de informática e os métodos utilizados nos processos de tratamento descritivo da informação (TDI) da área de Ciência da Informação consolidam a construção padronizada e consistente de metadados, proporcionando uma representação que garanta a unicidade e a recuperação dos recursos em diferentes ambientes informacionais digitais. Sendo assim, a hipótese de pesquisa é que as teorias, os princípios, os fundamentos, os métodos e as técnicas de Catalogação, que neste momento passam por uma fase de atualização perante as mudanças tecnológicas, constituem-se como metodologias que norteiam a construção padronizada e de representações informacionais unívocas nos padrões de metadados da área de Ciência da Informação. Neste sentido, a proposição desta pesquisa é analisar o referencial teórico que orienta a construção padronizada de representações nos padrões de metadados da Ciência da Informação. Por meio de uma metodologia de análise exploratória e descritiva da literatura disponível sobre o tema, apresenta-se como objetivo analisar as teorias, os princípios, os fundamentos e os instrumentos de representação da Catalogação e sua relação com os metadados, com o intuito de categorizar os instrumentos necessários para a construção de representações informacionais no âmbito do tratamento descritivo da informação. Com a realização dos objetivos, foi possível comprovar a tese de pesquisa e verificar que a integração estratégica entre as tecnologias de informática e os métodos utilizados nos processos de tratamento descritivo da informação (TDI) na Ciência da Informação consolidam a construção padronizada e consistente de metadados, proporcionando uma representação que garante a unicidade e a recuperação dos recursos em diferentes ambientes informacionais digitais. Como conclusão, ressalta-se que os metadados desenvolvidos e padronizados no domínio bibliográfico constituem-se como elementos essenciais do processo de TDI, não havendo necessidade de uso de outros padrões que não sejam os já desenvolvidos no domínio. Neste sentido, destaca-se a importância da interação entre teoria e prática nos processos de TDI, em especial no desenvolvimento de aplicações práticas dos instrumentos de representação desenvolvidos atualmente, para que possa ser aprimorado o tratamento descritivo da informação, diante dos avanços tecnológicos atuais.

Palavras-chave: Metadados. Tratamento Descritivo da Informação. Padrões de Metadados. Ambientes Informacionais Digitais. Metodologias de Padronização de Metadados.

ABSTRACT

The 21st century has been marked by related changes to information and technology, notably as regards the production and use of information in the computing technologies and in the development of information and communication technologies (ICTs) to disseminate them. Thus, the representation of information has become requirement in many areas of knowledge and to meet this need were developed metadata standards that diversify from simple structures, through a kind of intermediate standard to more complex structures description. The diversity of standards Information Science area is faced with the following impasse that constitutes a research problem: how to construct appropriate and standardized representations to ensure unity and efficient retrieval of information resources? This research which belongs to information and technology line has as theme the analysis of forms of representation of information resources, in view of the changes caused by advances in ICTs in the descriptive treatment of information (DTI) with the emergence of metadata. The focus is on analysis of the theories, principles, fundamentals and tools that guide of the standardized representations development in the Information Science field. The thesis hold for this research is that the strategic integration between computing technologies and methods used in descriptive treatment of information (DTI) processes of the Science Information field consolidate the consistent and standardized construction of metadata providing a representation which ensures the unity and resources retrieval in digital information environments different. Thus, the research hypothesis is that the theories, principles, fundamentals, methods and techniques of cataloging, which are currently going through technological changes, are methodologies that guide the standardized construction and univocal informational representations in metadata standards of the Information Science. Thus, the research hypothesis is that the theories, principles, fundamentals, methods and techniques of cataloging, which are currently going through technological changes, are methodologies that guide the standardized construction and univocal informational representations in metadata standards of the Information Science. The proposition of this research is to analyze the theoretical reference that guides the standardized representations construction in metadata standards of Information Science. Through an exploratory and descriptive analysis methodology of the available literature on the subject is presented to analyze the theories, principles, fundamentals and tools of representation of the Cataloging and its relationship with the metadata in aim to categorize necessary tools for the informational representations construction in descriptive treatment of information scope. With the achievement of the objectives was possible to prove the research thesis and verify that the strategic integration between the computing technologies with the methods used in descriptive treatment of information (DTI) processes in Information Science consolidate the consistent and standardized construction of metadata providing a representation which ensures the unity and resources retrieval in digital information environments different. In conclusion we highlighted that developed and standardized metadata in bibliographic domain constitute as essential elements of the DTI process, and there is no need to use other standards than those already developed in the domain. In this sense, we highlight the importance of the interaction between theory and practice in the DTI process, in particular in the development of practical applications of the instruments of representation currently developed, so it can be improved descriptive treatment of information before of current technological advances.

Keywords: Metadata. Descriptive treatment of information. Metadata Standards. Digital information environments. Metadata Standardization methodologies.

LISTA DE QUADROS

QUADRO 1: Sistematização da história da Catalogação em períodos	41
QUADRO 2: Dupla perspectiva do papel dos metadados	51
QUADRO 3: Tipologia de formatos de metadados	60
QUADRO 4: Elementos do Esquema de Metadados DC	65
QUADRO 5: Exemplo do campo Título e indicação de responsabilidade, do formato MARC 21	73
QUADRO 6: Esquemas de codificação e padrões correspondentes	92
QUADRO 7: Sistematização das contribuições para a teoria da catalogação.	104
QUADRO 8: Sistematização dos princípios e fundamentos norteadores do domínio bibliográfico	110
QUADRO 09: Categorização dos instrumentos de representação no TDI no domínio bibliográfico	117
QUADRO 10: Síntese geral dos componentes do TDI	118

SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 PROBLEMAS DE PESQUISA.....	14
1.2 HIPÓTESE, TESE E PROPOSIÇÃO DA PESQUISA.....	16
1.3 OBJETIVOS.....	17
1.3.1 OBJETIVO GERAL.....	18
1.3.2 OBJETIVOS ESPECÍFICOS.....	18
1.4 METODOLOGIA.....	19
1.5 JUSTIFICATIVA.....	21
1.6 ESTRUTURA DO TRABALHO.....	22
2 DA CATALOGAÇÃO AOS METADADOS: o tratamento descritivo da informação e as tecnologias de informática.....	25
2.1 PERÍODOS REMOTOS.....	26
2.2 PERÍODO TIPOGRÁFICO.....	28
2.3 PERÍODO TRADICIONAL.....	30
2.4 PERÍODO PRÉ-MECANIZADO.....	32
2.5 PERÍODO MECANIZADO.....	34
2.6 PERÍODO DE METADADOS.....	37
2.7 SISTEMATIZAÇÃO DA INTEGRAÇÃO ESTRATÉGICA DA TDI COM AS TECNOLOGIAS.....	40
3 METADADOS: revisando fundamentos, princípios, conceitos, tipologia, funções, características e aplicações.....	43
3.1 TIPOS, CARACTERÍSTICAS, FUNÇÕES, APLICAÇÕES E IMPORTÂNCIA DOS METADADOS.....	48
3.2 PRINCÍPIOS GERAIS.....	52
4 PADRÕES DE METADADOS.....	59
4.1 PADRÕES DE METADADOS PARA PROPÓSITOS GERAIS: formatos estruturados para o domínio da Web.....	62
4.2 PADRÕES DE METADADOS PARA PROPÓSITOS ESPECÍFICOS: formatos ricos para o domínio bibliográfico.....	68
5 ASPECTOS TECNOLÓGICOS E REPRESENTACIONAIS PARA A DETERMINAÇÃO PADRONIZADA DOS METADADOS.....	77
5.1 ASPECTOS TECNOLÓGICOS PARA DETERMINAÇÃO DOS METADADOS: requisitos funcionais, modelagem conceitual e esquemas de metadados.....	79
5.1.1 REQUISITOS FUNCIONAIS E MODELAGEM CONCEITUAL.....	79
5.1.2 ESQUEMAS DE METADADOS: conjunto de elementos e espaço do valor.....	85
5.1.2.1. CONJUNTO DE METADADOS (<i>ELEMENT SET</i>).....	87
5.1.2.2 ESPAÇO DO VALOR (<i>VALUE SPACES</i>).....	89
5.2 ASPECTOS REPRESENTACIONAIS PARA A DETERMINAÇÃO	

DOS METADADOS: esquemas de codificação.....	90
6 TEORIAS, PRINCÍPIOS, FUNDAMENTOS, MÉTODOS E INSTRUMENTOS DO TRATAMENTO DESCRITIVO DA INFORMAÇÃO NO DOMÍNIO BIBLIOGRÁFICO.....	97
6.1 CONTRIBUIÇÕES PARA UMA TEORIA DA CATALOGAÇÃO.....	99
6.1.1 PRINCÍPIOS DE PANIZZI.....	99
6.1.2 PRINCÍPIOS DE JEWETT.....	100
6.1.3 PRINCÍPIOS DE CUTTER.....	100
6.1.4 PRINCÍPIOS DE LUBETSKY.....	102
6.1.5 CONTRIBUIÇÕES DE GORMAN.....	103
6.1.6 CONSIDERAÇÕES GERAIS SOBRE AS CONTRIBUIÇÕES PARA UMA TEORIA DA CATALOGAÇÃO.....	104
6.2 PRINCÍPIOS E FUNDAMENTOS: Declaração dos Princípios Internacionais de Catalogação - Princípios de Paris.....	105
6.2.1 INSTRUMENTOS DE REPRESENTAÇÃO: modelos conceituais, esquemas de codificação, padrões de metadados.....	111
6.2.1.1 MODELO CONCEITUAL.....	112
6.2.1.2 ESQUEMAS DE CODIFICAÇÃO.....	114
6.2.1.3 PADRÃO DE METADADOS.....	116
6.3 TENDÊNCIAS ATUAIS NO MÉTODO DE TRATAMENTO DESCRITIVO DA INFORMAÇÃO NO DOMÍNIO BIBLIOGRÁFICO..	119
7 CONSIDERAÇÕES FINAIS.....	122
REFERÊNCIAS.....	127

1 INTRODUÇÃO

O rápido desenvolvimento das tecnologias de informática, em especial das Tecnologias de Informação e Comunicação – TICs – vem ocasionando profundas transformações não só no contexto sociocultural, mas também em diversas áreas do conhecimento científico. Neste cenário, marcado pelo surgimento diversificado de ambientes informacionais digitais, a informação adquire status cada vez mais importante, por ser considerada insumo para a viabilização de atividades e de processos de tomada de decisões, nos processos culturais, nos processos de ensino e aprendizagem, nos processos de aquisição e construção de conhecimento para pesquisa científica e tecnológica e em diversos outros setores ou atividades nos quais o conhecimento humano se faz necessário (MARCONDES, 2001).

Desse modo, uma das principais mudanças no início do século XXI está relacionada à informação e à tecnologia, mais especificamente, na produção e no uso da informação, nas tecnologias de informática e desenvolvimento das TICs utilizadas para difundir essas informações.

O crescimento exponencial das informações e o surgimento de ambientes informacionais em meio digital são consequências do desenvolvimento e importância que o item informação e TICs desempenham na sociedade atualmente. Contudo, essa realidade causa um aspecto problemático: a dificuldade de identificação, organização, busca, recuperação, acesso e uso das informações em ambientes digitais.

O valor da informação como insumo está relacionado com o seu potencial de orientar a realização de uma determinada atividade. Para que possa expressar seu valor e potencial, “[...] a informação relevante para um dado problema precisa estar disponível no tempo certo. De nada adianta a informação existir, se quem dela necessita não sabe da sua existência, ou se ela não puder ser encontrada”. (MARCONDES, 2001, p. 61). Nesse sentido, a rapidez no processamento, na busca, na recuperação, no acesso, na difusão e na aquisição da informação está intrinsecamente ligada ao uso das TICs. E isso se apresenta como um fator preponderante para que a informação possa expressar seu valor e ser utilizada da melhor forma.

Nesse contexto, a área de Ciência da Informação se depara com um ponto crítico a ser solucionado e que ao mesmo tempo se constitui na essência e no ponto de origem dessa ciência: a necessidade de tornar acessível e disponível o crescente volume de informações produzidas, bem como a necessidade de solucionar os problemas relacionados à recuperação dessas informações (SARACEVIC, 1996).

É importante esclarecer e contextualizar nesta pesquisa o sentido do termo acessível, pois embora seja utilizado há muito tempo na área de Ciência da Informação pode trazer certa ambiguidade em relação ao termo acessibilidade, que é definida por Ferreira (2004) como: “Condição de acesso aos serviços de informação, documentação e comunicação, por parte de portador de necessidades especiais”. De acordo com Ferreira (2004), o termo acessível pode ser definido como o “que se pode alcançar, obter ou possuir”, em outras palavras, encontrar, adquirir ou obter acesso aos recursos informacionais disponíveis em um ambiente informacional (IFLA, 2009).

Os problemas apontados por Saracevic (1996) são intrínsecos à área de Ciência da Informação e a solução encontrada está pautada nos métodos que desenvolve para resolvê-los. Pode-se dizer, então, que as metodologias advindas do caráter interdisciplinar da Ciência da Informação, bem como o uso intensivo de tecnologias de informática, caracterizam as soluções estudadas para estes problemas.

Não é objetivo deste trabalho tratar da epistemologia da Ciência da Informação, mas apenas apontar algumas características importantes para contextualização desta pesquisa. Assim, abordam-se aqui os conceitos técnicos abarcados pela Ciência da Informação, ou seja, os aspectos empíricos relacionados com os métodos de análise dos recursos informacionais, sendo um deles a catalogação (LE COADIC, 2004).

Os conceitos técnicos da Ciência da Informação referem-se à gestão e armazenamento de documentos em um sistema de informação e estão diretamente relacionados com os métodos utilizados por essa ciência para análise dos documentos. Esses métodos constituem-se na catalogação, na indexação, na elaboração de resumos, entre outros (LECOADIC, 2004). Dentre os métodos citados, a catalogação constitui-se como um processo de representação informacional que garante a identificação única do recurso informacional para fins de recuperação e, portanto, essencial para proporcionar a recuperação eficiente dos recursos nos variados tipos de sistemas de informação. Desenvolvida na área da Biblioteconomia, a catalogação é um método para construção de representações do qual a Ciência da Informação se utiliza de forma interdisciplinar.

Contudo, o aprimoramento das TICs e o decorrente surgimento de novos tipos de recursos e ambientes informacionais vêm colocando em discussão os métodos tradicionais de processamento e tratamento informacional, tradicionalmente conhecidos como catalogação, classificação e indexação, desenvolvidos e aplicados há muito tempo na área de Biblioteconomia (MÉNDEZ RODRÍGUEZ, 2002).

Muitas são as denominações utilizadas ao longo do tempo para designar o tratamento informacional. Para esta pesquisa serão utilizados os termos tratamento descritivo da informação (TDI) para designar a representação informacional de um recurso.

Deste modo, é importante esclarecer o significado desses termos para a presente pesquisa. O termo tratamento descritivo da informação (TDI) designa o processo de representação informacional do recurso, ou seja, individualização e caracterização de um item documentário, diferenciando-o do tratamento temático também realizado na área de Biblioteconomia.

Já o termo representação informacional é empregado com o objetivo de abarcar o resultado do processo do tratamento descritivo da informação (TDI), que seria o conjunto de características que individualizam um recurso informacional. Opta-se por esse termo, pois considera-se toda representação uma forma de descrição e, portanto, seria redundante o uso do termo representação descritiva.

E, por fim, o termo recurso ou recurso informacional refere-se ao documento, item informacional, a informação registrada, a informação como coisa ou ainda a obra intelectual (entidade) expressa sob diversas manifestações (tradicionais ou digitais).

Em relação às mudanças tecnológicas que vêm ocorrendo e seus impactos no tratamento descritivo da informação (TDI) e na recuperação da informação, podem-se destacar, de modo geral, dois aspectos: o primeiro refere-se à concepção dos sistemas de informação, exigindo um esforço e uma redefinição de conceitos e tarefas do profissional da informação em relação à gestão da informação, ocasionados pelo desenvolvimento e aplicação de tecnologias de informática. O segundo aspecto, intimamente ligado com o primeiro por ser decorrente da adoção das tecnologias de informática e da consolidação de ambientes informacionais digitais, refere-se às mudanças no tratamento descritivo da informação, principalmente com o desenvolvimento, aplicação e uso de metadados e padrões de metadados como nova forma para designar os elementos de representação dos recursos em novos tipos de ambientes informacionais (MÉNDEZ RODRÍGUEZ, 2002).

O desenvolvimento dos ambientes informacionais digitais exige o desenvolvimento concomitante de novas infraestruturas de informação, ou seja, além da infraestrutura

tecnológica é necessário o desenvolvimento de uma “[...] infoestrutura coerente que assegure não só o acesso, mas também a recuperação e o intercâmbio de informação. [...]” (MÉNDEZ RODRÍGUEZ, 2002, p. 20).

Nesse contexto, a Ciência da Informação se destaca como uma importante aliada na busca de soluções para esses problemas devido a sua natureza científica. De acordo com Saracevic (1996, p. 42), “[...] a Ciência da Informação está inexoravelmente ligada à tecnologia da informação. O imperativo tecnológico determina a CI [...]”, dessa forma, pode-se considerar que o imperativo tecnológico estará presente na maior parte dos processos e metodologias utilizadas por esta ciência.

A origem da Ciência da Informação está no crescimento exponencial das informações e a solução apontada está na aplicação de tecnologias de informação, atualmente as TICs. Dessa forma, disponibilizar as informações de modo que possam ser recuperadas passa a ser o desafio, que engloba tanto aspectos tecnológicos como aspectos de representação informacional. Contudo, toda solução origina problemas mais específicos e próprios devido ao constante desenvolvimento que enfrenta (SARACEVIC, 1996).

Dessa forma, vivencia-se atualmente uma revisão dos fundamentos no tratamento descritivo da informação (TDI), tendo como base o uso intensivo de tecnologias de informática que se consolidam como uma característica inerente a esse processo e com o conseqüente surgimento de novas formas de tratamento descritivo da informação, os metadados, que são considerados como atributos que representam um recurso informacional.

Portanto, acredita-se que os metadados se constituem como um elemento fundamental no processo de tratamento descritivo da informação, pois refletem a conjunção entre aspectos tecnológicos e de representação necessários aos novos tipos de recursos e ambientes informacionais surgidos, facilitando assim, uma recuperação mais eficiente em ambientes digitais.

1.1 PROBLEMAS DE PESQUISA

Neste cenário, destacam-se os metadados como elementos intrínsecos aos sistemas de informação e aos ambientes informacionais digitais, sendo considerados instrumentos fundamentais para promover a representação, a individualização, o intercâmbio, a interoperabilidade entre sistemas, o acesso e a recuperação de recursos informacionais.

Contudo, somente a aplicação adequada e normalizada de metadados irá assegurar a efetividade dos sistemas. Em outras palavras, uma representação adequada é considerada uma representação que corresponde de modo exato a seu objeto e, portanto, deve ser padronizada (FERREIRA, 2004). Quanto maior a padronização na construção e aplicação de metadados, maior será a economia dos sistemas, proporcionando uma recuperação com bons resultados.

A necessidade de representar as informações existe em diversas áreas do conhecimento. Assim, foram sendo criados vários padrões de metadados que variam desde estruturas simples, passando por um tipo de padrão intermediário, até padrões de estruturas mais complexas de descrição. O uso do padrão de metadados correspondente ao tipo específico de ambiente informacional, bem como a construção adequada e normalização das representações constituem-se no ponto-chave para a efetividade dos sistemas de informação em ambientes digitais.

Se os metadados se constituem como um novo paradigma do tratamento descritivo da informação por integrarem aspectos tecnológicos e de representação e se existem em uma variedade de tipos e níveis, a área de Ciência da Informação se depara com um impasse que se constitui como um *problema de pesquisa*: como construir representações adequadas e padronizadas que garantam a unicidade e a recuperação eficiente dos recursos informacionais?

Outras questões constituem-se como alvo deste estudo, contribuindo para especificar o problema de pesquisa:

- Tendo como base seu caráter interdisciplinar, a Ciência da Informação dispõe de uma teoria e uma metodologia para o tratamento descritivo da informação (TDI)?
- Os métodos atuais de tratamento descritivo da informação (TDI) são adequados para a construção padronizada de representações dos recursos informacionais?
- Qual a relação do tratamento descritivo da informação da área de Ciência da Informação com os metadados?
- As diferentes representações se distinguem entre si pelo tipo de entidade que busca representar, pelo tipo de necessidade a que busca atender e pelo tipo de ambiente informacional na qual está inserida. Dessa forma, quais são as teorias, os princípios, os fundamentos, os métodos e os instrumentos que norteiam a construção de representações? Quais requisitos a representação deve seguir? Como ocorre essa representação no domínio bibliográfico atualmente? Como ocorre essa representação em um domínio emergente como a Web?

Deste modo, o *tema* de pesquisa consiste na análise das Formas de Representação dos recursos informacionais, tendo em vista as mudanças causadas pelos avanços das TICs no tratamento descritivo da informação (TDI) com o surgimento dos metadados. O *foco* desta pesquisa constitui-se na análise das teorias, dos princípios, dos fundamentos e dos instrumentos que norteiam a construção de representações padronizadas na área de Ciência da Informação.

1.2 HIPÓTESE, TESE E PROPOSIÇÃO DA PESQUISA

De modo geral, pode-se dizer que a área de Ciência da Informação busca investigar a geração, a organização, o processamento, a recuperação e o uso da informação, em diversos ambientes informacionais, incluindo o digital. Neste âmbito, o trabalho informacional de desenvolvimento e sistematização de representações dos recursos apresenta-se como primordial no intuito de proporcionar a efetividade dos sistemas nesses ambientes. Ressalta-se então que a construção padronizada de representações torna-se essencial.

Deste modo, a *tese* levantada para esta pesquisa é de que a integração estratégica entre as tecnologias de informática e os métodos utilizados nos processos de tratamento descritivo da informação (TDI) da área de Ciência da Informação consolidam a construção padronizada e consistente de metadados, proporcionando uma representação que garanta a unicidade e a recuperação dos recursos em diferentes ambientes informacionais digitais.

Sendo assim, a *hipótese* de pesquisa é que as teorias, os princípios, os fundamentos, os métodos e as técnicas de catalogação, que neste momento passam por uma fase de atualização perante as mudanças tecnológicas, constituem-se como metodologias que norteiam a construção padronizada e de representações informacionais unívocas¹ nos padrões de metadados da área de Ciência da Informação.

É importante destacar que o tratamento descritivo da informação da área de Ciência da Informação necessita de explicitação sobre seus fundamentos em relação aos seus aspectos representacionais e exige aperfeiçoamentos constantes para que possa estar em consonância com os avanços tecnológicos. Neste sentido, a *proposição* desta pesquisa é analisar o

¹ Que só comporta uma forma de interpretação (FERREIRA, 2004),

referencial teórico que orienta a construção padronizada de representações nos padrões de metadados da Ciência da Informação.

1.3 OBJETIVOS

Os metadados são inerentes aos sistemas de informação, porém, existem em uma variedade de tipos e níveis, criados para atender a necessidades distintas de representação em diferentes domínios de conhecimento.

Para melhor entendimento é preciso, primeiramente, definir o que vem a ser domínio. O termo é definido por Ferreira (2004) como “âmbito de uma arte ou de uma ciência [...]”, ou pode ser definido ainda como o “conjunto de assuntos de que trata uma ciência ou arte: área” (SANTOS, 2001). Deste modo, o termo domínio é considerado nesta pesquisa como o âmbito ou os assuntos abarcados em uma ciência, área, disciplina e comunidade do conhecimento.

Assim, o domínio bibliográfico destacado nesta pesquisa é considerado sinônimo de universo bibliográfico, definido pela IFLA (2009) como o “domínio relacionado a coleções de bibliotecas, arquivos, museus e outras comunidades de informação” (IFLA, 2009). Já o domínio Web pode ser entendido como o universo relacionado aos diferentes tipos de ambientes informacionais digitais, bem como os variados tipos de recursos disponibilizados por distintas áreas do conhecimento.

Embora o termo metadados seja uma denominação atual no processo de TDI, é importante considerar que os metadados não se constituem como algo novo para as áreas de Biblioteconomia e de Ciência da Informação, que vêm produzindo e padronizando metadados ao longo do tempo. Recentemente profissionais de diferentes áreas têm buscado criar outros métodos de tratamento descritivo da informação, sem saber que isso já foi feito, gerando uma variedade de padrões que muitas vezes não atende satisfatoriamente às necessidades informacionais já estabelecidas e padronizadas por outros atributos descritivos definidos nos processos de TDI da Ciência da Informação (MILSTEAD; FELDMAN, 1999).

Desta forma, para a verificação da tese e da hipótese propostas, esta pesquisa apresenta os seguintes objetivos:

1.3.1 OBJETIVO GERAL

O *objetivo geral* desta pesquisa é analisar as teorias, os princípios, os fundamentos e os instrumentos de representação da Catalogação e sua relação com os metadados, com o intuito de categorizar os instrumentos necessários para a construção de representações informacionais no âmbito do tratamento descritivo da informação.

1.3.2 OBJETIVOS ESPECÍFICOS

Para que o resultado dessa pesquisa seja alcançado, de acordo com o objetivo geral proposto, foram estabelecidos os seguintes *objetivos específicos*:

- Apresentar, na área de tratamento descritivo da informação (TDI) na Ciência da Informação, o referencial que orienta a definição dos dados referenciais (metadados) que representam o recurso informacional e as estruturas padronizadas de descrição (conjunto de dados-atributos ou metadados);
- Analisar a integração estratégica das tecnologias de informática com o tratamento descritivo da informação (TDI) desenvolvido na disciplina de Catalogação e sua relação com os metadados para a construção de representações informacionais;
- Atualizar o conceito de metadados e padrões de metadados para a área de Ciência da Informação, por meio de uma revisão sistematizada;
- Analisar os tipos, os níveis e a atribuição de metadados no domínio da Ciência da Informação e sua relação com o domínio da Web;
- Analisar os instrumentos utilizados no tratamento descritivo da informação na Catalogação atualmente e sua relação com as tecnologias de informática;
- Com base na análise das teorias, dos princípios, dos fundamentos da Catalogação e a relação com os metadados, propor uma categorização dos instrumentos utilizados para a padronização de representações em padrões de metadados específicos do domínio bibliográfico, conforme pode ser observado nas sistematizações apresentadas nos capítulos 5 e 6.

1.4 METODOLOGIA

Para o desenvolvimento da pesquisa, a *metodologia* utilizada consistiu em uma análise exploratória e descritiva da literatura disponível sobre o tema proposto, o que permitiu a construção de um conhecimento teórico sobre as teorias, os princípios, os fundamentos e os instrumentos que orientam a construção de representações na área de Ciência da Informação, auxiliando, assim, na compreensão do problema proposto e na construção dos resultados esperados.

O caráter exploratório da pesquisa, segundo Cervo e Bervian (2003), consiste na busca de informações sobre o assunto a ser estudado, considerando os diversos aspectos de um problema a fim de solucioná-lo. Segundo Köcher (2002, p. 126), estudos exploratórios desencadeiam “[...] um processo de investigação que identifique a natureza do fenômeno e aponte as características essenciais das variáveis que se quer estudar”. Além disso, o caráter descritivo da pesquisa permitirá descrever as características do objeto de estudo em questão, conforme especificadas na tese, hipótese, proposição e objetivos (COSTA, COSTA, 2001).

O *universo da pesquisa* se constituiu nas teorias, nos princípios, nos fundamentos e nos instrumentos da disciplina de Catalogação; nas teorias sobre os metadados e padrões de metadados para o tratamento descritivo da informação, bem como na integração estratégica destes com as tecnologias de informática. O estudo e análise do universo da pesquisa possibilitou um entendimento melhor sobre as soluções apontadas para resolver o problema de pesquisa.

A partir do estudo exploratório e descritivo da literatura, foi possível concretizar os objetivos propostos sobre o universo da pesquisa e chegar a algumas soluções e considerações sobre o problema de pesquisa.

Deste modo, os *procedimentos metodológicos* estão divididos nas seguintes *etapas de realização da pesquisa*, descritas a seguir:

1ª Etapa - Levantamento bibliográfico e seleção do material obtido: busca de informações pertinentes ao tema, com o intuito de criar o embasamento teórico ao estudo proposto.

O levantamento bibliográfico foi efetuado em nível nacional e internacional e em fontes bibliográficas da área de estudo, divididas nos seguintes tipos:

- a) Fontes primárias: livros, periódicos, anais de congressos, dissertações e/ou teses, documentos eletrônicos da Internet e outros documentos congêneres.
- b) Fontes secundárias: base de dados nacionais e internacionais: PROBE, Web of Science, LISA, Periódicos Capes, Current Contents, Scielo, Science Direct, entre outras fontes de informação julgadas pertinentes no decorrer da pesquisa, tais como sites especializados da área que continham publicações relevantes sobre o tema (ex.: IFLA);

Quanto ao material obtido, a seleção foi realizada a partir dos seguintes critérios:

- a) Pertinência ao tema escolhido;
- b) Documentos em português, inglês, francês e espanhol;
- c) Período limitado aos últimos dez anos apenas como abordagem inicial; não houve limitação cronológica para referências bibliográficas citadas nos documentos identificados.

2ª Etapa – Leitura, interpretação, análise e sistematização das informações: realizada nos documentos selecionados, a fim de se oportunizar o desenvolvimento da base teórica para a discussão dos diferentes pontos de vista identificados na literatura sobre o tema proposto, possibilitando, assim, criar subsídios para a elaboração da pesquisa.

3ª Etapa – Análise e estabelecimento das características fundamentais extraídas da literatura: verificação e análise das principais características encontradas na literatura sobre o tema para elucidação do problema de pesquisa, criando, assim, a base teórica para elaboração (redação) da pesquisa.

4ª Etapa – Sistematização do estudo exploratório: com o intuito de identificar e analisar as relações existentes entre as teorias, os princípios, os fundamentos e os instrumentos de representação da Catalogação e sua relação com os metadados. É na sistematização do estudo exploratório que se elucida o problema de pesquisa por meio dos objetivos propostos a fim de atingir os resultados esperados da pesquisa, isto é, categorizar os instrumentos necessários para a construção de representações informacionais unívocas no domínio bibliográfico.

5ª Etapa – Elaboração e redação para o exame de qualificação: apresentação, à banca examinadora, das considerações preliminares sobre o estudo proposto.

6ª Etapa – Elaboração e redação final da pesquisa: após a orientação e sugestões da banca de qualificação, desenvolvimento das propostas finais da pesquisa com o intuito de divulgação à comunidade científica dos resultados obtidos com o desenvolvimento do estudo em questão.

1.5 JUSTIFICATIVA

São incontestáveis as mudanças pelas quais nossa sociedade vem passando devido à evolução das tecnologias de informática, principalmente com o desenvolvimento e a implementação das TICs. Essas transformações ganham uma dimensão cada vez maior, principalmente por integrarem vários modos de comunicação em uma multiplicidade de ambientes informacionais digitais.

Por sua vez, a informação vem adquirindo um valor crescente nesses ambientes e promover de modo rápido e fácil o processamento, a busca, a recuperação, o acesso, a difusão e a aquisição da informação constitui-se como fator primordial no contexto sociocultural na atualidade. A literatura aponta que a solução está pautada na aplicação e no uso de padrões de metadados. Contudo, em um estudo anteriormente desenvolvido, foi possível verificar que o problema não está somente na falta de uso dos padrões de metadados, mas principalmente na falta de aplicação apropriada dos instrumentos necessários para a construção de representações, pois são responsáveis por garantir a unicidade do recurso informacional.

Embora existam muitos estudos sobre os metadados, ainda é comum encontrar na literatura alguns equívocos sobre suas características, principalmente ao considerá-los elementos de representação desenvolvidos exclusivamente para o domínio Web, fato que não se comprova quando observamos o desenvolvimento de metadados no domínio bibliográfico, por exemplo. Esse equívoco ocorre devido ao desconhecimento da variedade de tipos e níveis de padrões de metadados e aos diferentes domínios de aplicação e uso, o que gera uma diversidade de tipos de representação.

O desenvolvimento da pesquisa torna-se de significativa importância, pois procura elucidar as características dos metadados, sua relação com o tratamento descritivo da informação (TDI) na área de Ciência da Informação e sua relação com a Catalogação. O intuito é demonstrar que o processo de tratamento descritivo da informação (TDI), realizado por meio das teorias, dos princípios, dos fundamentos e dos instrumentos da Catalogação, vem passando por uma reavaliação diante dos avanços das tecnologias de informática, com o

objetivo de se adequar à nova realidade dos ambientes informacionais digitais. Diante disso, torna-se importante destacar que o TDI passa a ser fundamentalmente baseado nas tecnologias de informática e que as teorias, os princípios, os fundamentos e os instrumentos da Catalogação, constituem-se como metodologias que orientam a construção unívoca e padronizada dos metadados e suas representações no domínio bibliográfico.

Dessa forma, o tema estudado constitui-se como relevante para a área de Ciência da Informação por analisar e elucidar os aspectos atuais relacionados ao uso estratégico das tecnologias de informática e à construção padronizada de representações dos recursos informacionais, fator importante a ser definido no processo de tratamento descritivo da informação.

Como relevância social, a realização desta pesquisa se justifica por mostrar à comunidade científica e ao público em geral as contribuições que a área de Ciência da Informação pode oferecer no âmbito da representação informacional e conseqüentemente na recuperação da informação, com o uso de metadados específicos e cuidadosamente construídos ao longo da história de tratamento descritivo da informação e consolidados no domínio bibliográfico.

Por esse motivo, acredita-se que uma das maiores contribuições desta pesquisa para a linha de pesquisa Informação e Tecnologia, do Programa de Pós-Graduação em Ciência da Informação, está em destacar as metodologias que orientam a construção padronizada de representações das informações, tendo como base os avanços tecnológicos, mostrando a integração estratégica entre tecnologias de informática e o tratamento descritivo da informação nos processos de representação informacional, facilitando assim o processamento, a busca, a recuperação, o acesso, a difusão e a aquisição da informação em ambientes informacionais digitais.

1.6 ESTRUTURA DO TRABALHO

Quanto à organização estrutural do trabalho, além do presente capítulo que aborda questões iniciais como o problema da pesquisa, tese, hipótese, justificativa, objetivos e metodologia, apresentam-se os seguintes capítulos conforme descritos a seguir:

- CAPÍTULO 2 - DA CATALOGAÇÃO AOS METADADOS: o tratamento descritivo da informação e as tecnologias de informática: abordam-se as principais questões relacionadas ao desenvolvimento do TDI no domínio bibliográfico, destacando a integração estratégica com as tecnologias, em especial as tecnologias de informática;
- CAPÍTULO 3 - METADADOS: revisando fundamentos, princípios, conceitos, tipologia, funções, características e aplicações: trata-se de uma revisão de fundamentos e conceitos, destacando em especial os princípios, as tipologias, as funções, as características e as aplicações dos metadados, a fim de contextualizá-los no cenário atual e posteriormente relacioná-los com a questão do TDI;
- CAPÍTULO 4 - PADRÕES DE METADADOS: abordam-se os aspectos referentes aos padrões de metadados, seus fundamentos, princípios, seus tipos e características, além dos aspectos principais de duas categorias de padrões: padrões de metadados para propósitos gerais (domínio da Web) e os padrões de metadados para propósitos específicos (domínio bibliográfico);
- CAPÍTULO 5 - ASPECTOS TECNOLÓGICOS E REPRESENTACIONAIS PARA A DETERMINAÇÃO PADRONIZADA DOS METADADOS: aborda os aspectos tecnológicos e representacionais necessários para a padronização dos metadados e conseqüentemente de representações;
- CAPÍTULO 6 - TEORIAS, PRINCÍPIOS, FUNDAMENTOS, MÉTODOS E INSTRUMENTOS DO TRATAMENTO DESCRITIVO DA INFORMAÇÃO NO DOMÍNIO BIBLIOGRÁFICO: apresentação do estudo exploratório que consiste na análise das teorias, dos princípios, dos fundamentos do TDI, bem como a categorização dos instrumentos de representação do universo bibliográfico e as tendências atuais no método de TDI no domínio bibliográfico.
- CAPÍTULO 7 – CONSIDERAÇÕES FINAIS: elaborada com base na análise do estudo exploratório, apresenta as considerações para a comprovação da tese;

- REFERÊNCIAS - Por fim, estão apresentadas as referências utilizadas no desenvolvimento desta pesquisa.

2 DA CATALOGAÇÃO AOS METADADOS: o tratamento descritivo da informação e as tecnologias de informática

A preocupação com os registros do conhecimento e com os métodos de sistematização dessas informações para a busca e recuperação sempre foi constante entre os profissionais envolvidos com a informação. Em todas as épocas houve uma busca de aprimoramento de técnicas para o tratamento descritivo da informação, no sentido de associá-las ao uso de alguma tecnologia que tornasse ágil e fácil esse processo.

Nesse sentido, pode-se considerar que a disciplina de Catalogação nasceu com a necessidade de estabelecimento de regras para a construção dos catálogos. Porém, foi sendo desenvolvida e aprimorada, tornando-se não só uma técnica para a construção de catálogos, mas principalmente uma metodologia para o processamento e tratamento descritivo da informação. Ao longo do tempo utilizou-se das tecnologias disponíveis em cada época como forma de aprimorar o processo de representação, com o intuito de facilitar a recuperação e disseminação dos recursos informacionais (PEREIRA, SANTOS, 1998).

Embora o histórico dos catálogos e da catalogação ressalte aspectos importantes sobre o tratamento descritivo da informação, não é objetivo deste capítulo abordá-lo de forma detalhada e exaustiva, pois isso já foi realizado por diversos autores, como Barbosa (1978); Mey (1995); Chan (2007); Strout (1956) abordado nos estudos de Mey e Silveira (2009).

O objetivo do capítulo é analisar os aspectos principais que consolidaram o tratamento descritivo da informação e sua integração estratégica com as tecnologias, em especial as tecnologias de informática. Em outras palavras, resgatam-se os aspectos e princípios que consolidaram as metodologias de construção padronizada de representações e sua relação com as tecnologias ao longo do tempo, até chegar ao uso de tecnologias de informática, principalmente até o desenvolvimento dos metadados, que refletem nitidamente a integração estratégica entre representação e tecnologias de informática.

Para tanto, este capítulo apresentará um breve histórico dos catálogos e da catalogação, até chegar ao momento atual, marcado pelo desenvolvimento e aplicação dos metadados e aprimoramento dos princípios, os fundamentos e os instrumentos da Catalogação.

Para melhor entendimento, o capítulo será dividido em alguns períodos históricos tendo como base os estudos de Barbosa (1978) e Mey e Silveira (2009); contudo, algumas

designações foram alteradas para melhor refletir os aspectos históricos, e outras foram criadas para determinar os novos acontecimentos não abrangidos nas designações originais, tal como o surgimento dos metadados. Cada um dos períodos apresenta uma peculiaridade; assim, destacam-se aqui os aspectos mais importantes e sua relação com as tecnologias.

2.1 PERÍODOS REMOTOS

A tentativa de se estabelecer sistematizações dos recursos informacionais sempre existiu em cada época, bem como o uso de tecnologias vigentes. (PEREIRA, SANTOS, 1998). Os Períodos Remotos compreendem o período da Antiguidade e da Idade Média e se caracterizam por: iniciativas incipientes em relação à construção dos catálogos e da catalogação; apresentar acervos específicos e restritos; usuários especializados; bem como o início da integração estratégica com as tecnologias vigentes.

Durante este período, a forma mais comum de organização dos recursos informacionais foram as listas de obras, parecidas com inventários das coleções. Embora também fossem utilizadas para encontrar os materiais no acervo, não é possível afirmar que elas se constituíam propriamente como um catálogo, pois eram elaboradas apenas para se ter o conhecimento das obras existentes no acervo de uma biblioteca (MEY, SILVEIRA, 2009). Além disso, as sistematizações eram feitas de modo arbitrário, ou seja, sem a construção padronizada de regras que norteassem a descrição dos materiais.

Conforme se observa nos estudos de Mey e Silveira (2009) a história dos catálogos e da catalogação na Antiguidade se caracteriza pelas primeiras iniciativas na sistematização de uma organização para a construção de catálogos e está diretamente relacionada à história das bibliotecas e à evolução dos suportes de informação.

De acordo com as autoras, a biblioteca mais antiga que se conhece até o momento (terceiro milênio antes de Cristo, localizada em Elba, atual Síria) possuía materiais organizados em estantes e apresentava um resumo dos documentos em tabletes de argila ou tábulas (MEY, SILVEIRA, 2009).

Escavações no Egito encontraram tábulas datadas de 1400 aC, referindo-se a títulos de obras. Entretanto, conforme destacam as autoras Mey e Silveira (2009, p. 60),

Datam de 1300 aC as tábulas com as primeiras informações bibliográficas de descrição física, descobertas em escavações hititas (atual Turquia). Essas

tábulas identificavam o número da tábula em uma série, o título e, muitas vezes o escriba.

Na biblioteca de Assurbanípal (ou biblioteca de Nínive), foram encontrados tabletes de argila com as seguintes informações descritas: título, número da tábula ou volume, as primeiras palavras da tábula seguinte, o nome do possuidor original, o nome do escriba e um selo de propriedade. De acordo com Mey e Silveira (2009, p. 60) “presume-se haver, nesta época, um embrião do catálogo”.

Destacam-se também neste período algumas iniciativas elaboradas por Calímaco na biblioteca de Alexandria. O “catálogo” de Calímaco, denominado *pinakoi* (que significa em grego estante, mesa ou tábua) consistia em uma organização dentro de grandes assuntos, listando os materiais existentes nas estantes e, provavelmente, nas mesas onde se deixavam os textos. Os materiais possuíam no verso pequenas etiquetas que identificavam a autor e título das obras. Calímaco contribuiu também para a criação de inúmeras obras de referência (bibliografias temáticas) (MEY, SILVEIRA, 2009).

Foi durante a Antiguidade que se estabeleceu um importante princípio para a sistematização de informações: “Os gregos foram responsáveis pela introdução do conceito de autor de uma obra como ponto de acesso a ela, princípio que permanece até nossos dias” (MEY, SILVEIRA, 2009, p. 62).

Na Idade Média, durante muitos séculos, “[...] os mosteiros foram os únicos preservadores, copistas e catalogadores de livros [...]” (MEY, SILVEIRA, 2009, p. 63). Nesta época, os catálogos ainda se caracterizavam como listas de obras ou inventários do acervo.

Uma das primeiras listas de obras de bibliotecas medievais (surgidas em meados do século VIII) apresentava informações como título e às vezes nome do autor, sem uma ordem visível. A partir disso, várias outras listas foram feitas, cada uma apresentando uma característica de organização e sistematização das informações. A biblioteca de Richenau, na Alemanha, por exemplo, fez uma compilação de vários catálogos que indicavam as obras contidas em cada volume e o número dos volumes; no mosteiro de Saint Requier, na França, o catálogo era organizado por autor sem uma ordem aparente, registrava o conteúdo dos volumes e o número dos volumes de uma obra. Durante algum tempo, os catálogos eram feitos sem grandes alterações. “Strout considera que todas as listas bibliográficas do período eram, na verdade, inventários.” (MEY, SILVEIRA, 2009, p. 64).

Pode-se considerar como um dos primeiros catálogos a lista do convento de Sanit Martin, em Dover, em 1389:

Dividia-se em três seções. A primeira, organizada pelo número de localização do volume na estante, incluía um título breve, o número da página do documento em que o número de localização foi registrado, as primeiras palavras do texto nesta página, o número de páginas do documento e o número de obras contidas no volume. A segunda, também organizada pelo número de localização, registrava o conteúdo de cada volume, com a paginação e as palavras iniciais de cada obra. A terceira é um marco na catalogação: incluía análise das partes (entradas analíticas) e uma lista alfabética, às vezes de autor, outras de título e autor e outras, ainda de palavras genéricas, como 'livro', 'parte' ou 'códice'. (MEY, SILVEIRA, 2009, p. 65).

Conforme pode ser observado nesse breve relato, observa-se que os catálogos constituíam-se como iniciativas incipientes na organização das bibliotecas durante os Períodos Remotos. De modo geral, o estabelecimento desses pré-catálogos estava vinculado à forma de organização física das bibliotecas e às características dos tipos de recursos informacionais. Portanto, não havia uma forma de organização e sistematização unificada das informações, cada pré-catálogo apresentava representações distintas. Além disso, dependiam das tecnologias vigentes em cada época, ou seja, dos diferentes tipos de suportes utilizados para registrar as informações (tabletes de argila, madeira, papiro, pergaminho, papel).

2.2 PERÍODO TIPOGRÁFICO

O período nomeado como tipográfico compreende os séculos XV a XVIII e se destaca tecnologicamente pela invenção da imprensa.

O advento da tipografia constituiu-se como um grande avanço tecnológico para a época e causou grande impacto na sociedade, que já passava por profundas mudanças sociais. O monopólio da informação começou a ser questionado, a informação passou a ser considerada sinônimo de desenvolvimento social e de igualdade; o surgimento de novos públicos leitores alterou as estruturas de circulação e comunicação da informação, novos tipos de recursos informacionais impressos começaram a surgir, houve um crescente aumento do número de publicações na época e surgiram as primeiras bibliotecas públicas, que passaram a ser consideradas essenciais para a democratização do conhecimento. (CHARTIER, 1998; WITTMANN, 1999; LYONS, 1999; MANGUEL, 2001; MARTINS, 2001).

O crescimento do número de publicações impressas acarretou uma preocupação com a normalização das publicações, surgiu a folha de rosto nos livros e a autoria começou a ser indicada (LITTON, 1975). Os recursos informacionais começaram a ser publicados de forma

normalizada, fato que iria auxiliar posteriormente o estabelecimento de representações com maior precisão nos catálogos.

No século XV surgiram pela primeira vez as remissivas, que apesar da forma primitiva, eram as primeiras iniciativas de registros remetendo a outros registros ou obras.

No início do século XVI, surgiu um catálogo classificado incluindo também um índice alfabético de autor. Em 1564, apareceu o primeiro catálogo de livreiros. Nos séculos XVII e XVIII, várias outras iniciativas de organização do acervo foram desenvolvidas em diferentes bibliotecas, dando origem às primeiras iniciativas de sistematização de regras para a construção dos catálogos.

Porém, de modo geral, conforme aponta Machado (2003, p. 42), pode-se considerar que

Os primeiros catálogos e bibliografias são puramente listas inventariais e não instrumentos bibliográficos. Nas bibliografias, a ênfase é dada aos autores e não aos livros, são biobibliografias; nos catálogos, únicos tipos de listas bibliográficas, a caracterização dá-se pela técnica pouco elaborada, falta de arranjo e transcrição sucinta e pouco precisa dos títulos (MELO, 1981; PINTO, 1987). Para melhor identificação das obras, que crescem quantitativamente com a invenção da imprensa, os títulos começam a ser utilizados e os acervos das bibliotecas e livrarias passam a exigir uma organização mais criteriosa.

Os catálogos em livros refletiam o avanço tecnológico da época, ocasionado pela imprensa. Mais do que isso, constituíam-se como “bibliotecas sem muros” e eram considerados instrumentos importantes para a democratização do acesso ao conhecimento impresso, embora ainda tivessem a característica de inventário de coleções. Seu formato em livro permitia maior portabilidade e, portanto, maior facilidade para a disseminação e troca de informações (CHARTIER, 1998).

Pode-se considerar a tecnologia proporcionada pela invenção da tipografia como o primeiro uso estratégico de tecnologias na construção dos catálogos, porém é importante ressaltar que na época não existiam ainda regras consolidadas de catalogação.

As primeiras iniciativas para a construção de um código de catalogação tiveram início somente no século XVIII. Durante a Revolução Francesa, várias bibliotecas foram confiscadas e transformadas, posteriormente, em bibliotecas públicas, exigindo a construção do primeiro código de catalogação em Paris (MEY, SILVEIRA, 2009).

Com os avanços em relação ao estabelecimento de regras, mudou o conceito de catálogo como inventário, e passou a ser considerado como lista para encontrar as obras. Também data do século XVIII o aparecimento dos catálogos em ficha, formato que perdura

até os dias de hoje e que possibilitou, na época, uma portabilidade ainda maior no intercâmbio dos registros bibliográficos (MEY, 1995; MEY, SILVEIRA, 2009).

Conforme destaca Barbosa (1978), as regras elaboradas para a redação das fichas advinham da necessidade de construção uniforme dos catálogos, portanto, não tinham a finalidade de se constituírem um código e não foram, de início, determinadas por bibliotecários.

A partir do século XVI, principalmente com o advento das bibliotecas públicas, haviam começado as primeiras preocupações com a normalização de algumas regras para a construção dos catálogos. Mas somente a partir do século XIX ocorrem os grandes estudos e trabalhos na tentativa de se estabelecer a construção de regras, padronizadas por princípios.

De modo geral, pode-se considerar neste período (séculos de XV a XVIII) que a tecnologia predominante era a tipografia, contribuindo para a construção de catálogos em forma de livros e posteriormente em fichas. Embora pareçam incipientes, as tecnologias na época trouxeram significativos avanços, pois possibilitaram a divulgação e o acesso amplo aos registros do conhecimento em diversos lugares. Além disso, os avanços proporcionados pela imprensa na construção dos catálogos predominam até os dias de hoje, obviamente em um novo contexto de tratamento descritivo da informação. Somente na metade do século XX é que começaram a ser utilizadas as primeiras tecnologias de informática como aliadas do processo de catalogação e da construção de catálogos.

2.3 PERÍODO TRADICIONAL

Considerado por Barbosa (1978) o período mais longo da história da catalogação, é demarcado pelos estudos de Antony Panizzi até a realização da Conferência de Paris (período de 1841 até 1961, final do século XIX e mais da metade do século XX).

Foi neste período que se iniciou grande parte dos estudos e desenvolvimentos de trabalhos importantes na catalogação, que influenciam as práticas do tratamento descritivo até os dias atuais (MEY, SILVEIRA, 2009).

De acordo com Chan (2007), muitos indivíduos e organizações contribuíram para o desenvolvimento de padrões e códigos visando a uma padronização na descrição bibliográfica. De modo geral, os esforços se converteram inicialmente para o estabelecimento de regras mais consistentes, que se converteram na construção de códigos de catalogação

(conjunto de regras para a descrição). Porém, para que isso não se constituísse em novas regras sem fundamentos, foram desenvolvidos, concomitantemente, princípios que nortearam construção das descrições, constituindo-se posteriormente, em teorias e princípios de catalogação.

Desta forma, conforme aponta Chan (2007), iniciaram-se neste período duas linhas de desenvolvimento na catalogação e que ocorreram concomitantemente: uma refere-se ao desenvolvimento e aprimoramento das regras de descrição e códigos de catalogação, a outra refere-se ao início do estabelecimento de teorias, princípios e fundamentos para a catalogação.

Em relação à primeira linha, destaca-se a consolidação de vários códigos de catalogação; o maior destaque vai para o código da ALA (American Library Association), as Instruções Prussianas e o Código da Vaticana (BARBOSA, 1978). Vários outros foram criados e muitas revisões foram feitas na tentativa de se estabelecer princípios e uma maior padronização das regras. Porém, não é objetivo aqui traçar as mudanças e alterações dos códigos criados, mas apenas destacar os códigos mais importantes do período.

Em relação à segunda linha, destaca-se a contribuição dos seguintes teóricos: Antony Panizzi, Charles C. Jewett, Charles Ami Cutter e Seymour Lubetsky. De modo geral, as contribuições desses teóricos referem-se ao estabelecimento de regras de catalogação, à revisão e aprimoramento das regras de catalogação da época e, principalmente, ao estabelecimento de algumas teorias e princípios para o processo de catalogação.

O impacto nos avanços tecnológicos após a Segunda Guerra Mundial deu origem a uma variedade de novos tipos de recursos informacionais, causando uma significativa mudança nos serviços de processamento de informação. Os catalogadores começaram a sentir necessidade de um código mais racional, pois gastavam muito tempo na elaboração das fichas. Essa necessidade resultou na Conferência de Paris, em 1961 (BARBOSA, 1978).

Após uma análise exaustiva dos códigos da época, elaborada por Seymour Lubetsky, verificou-se a existência de muitas regras desnecessárias, omissão de dados essenciais e algumas inconsistências nas regras em uso. A Conferência de Paris ou a Conferência Internacional sobre os Princípios de Catalogação foi realizada com o intuito de uniformizar as regras e estabelecer princípios para a Catalogação. A partir dela, vários códigos passaram por atualizações e vários outros códigos foram desenvolvidos (BARBOSA, 1978).

Com o desenvolvimento das teorias, dos princípios e dos fundamentos, foi possível a construção de regras e códigos mais consistentes e padronizados, fato que se consolidou ainda mais no século XX, diante da necessidade de padronização para automação dos processos de catalogação e que será abordado nos períodos a seguir.

Neste período as tecnologias proporcionadas pela imprensa ainda predominavam, embora já houvesse o desenvolvimento e a utilização de algumas tecnologias mecânicas que proporcionavam a automação de alguns processos. Um exemplo são os cartões perfurados criados por Herman Hollerith e utilizados no início do século XX nos Estados Unidos para registrar dados do censo: “[...] sistemas mecânicos foram usados para processar os cartões e tabular os resultados. Mais tarde, os cartões perfurados passaram a ser amplamente usados como um meio de inserir dados em computadores.” (SILBERSCHATZ; KORTH; SUDARSHAN, 2006, p. 19).

Entre os anos de 1950 até 1960, as tecnologias que se destacaram no processamento da informação também estavam relacionadas às tecnologias de armazenamento. Foram desenvolvidas as fitas magnéticas, e seu processamento “[...] consistia em ler dados em uma ou mais fitas e escrevê-los em uma nova fita. Os dados também podiam ser inseridos por decks de cartão perfurado e enviados para saída de impressoras. [...]”. (SILBERSCHATZ; KORTH; SUDARSHAN, 2006, p. 19).

2.4 PERÍODO PRÉ-MECANIZADO

Este período é demarcado pela Conferência de Paris até a realização da RIEC (Reunião Internacional de Especialistas em Catalogação), compreendendo o período de 1961 até 1969 (BARBOSA, 1978).

Embora seja um período curto quando comparado com o anterior, apresentou significativos avanços no que se refere ao estabelecimento de princípios e ao surgimento de formas automatizadas no processo de catalogação.

Os “Princípios de Paris”, documento que resultou da Conferência Internacional sobre Princípios de Catalogação (Conference on Cataloguing Principles) em Paris- França, foi estabelecido na tentativa de criar princípios que norteassem a construção lógica de regras de catalogação para comporem um código consistente. Para tanto, definia objetivos baseados nos teóricos da área e na análise das regras existentes que se constituem como fundamentos da catalogação (GAVIN, 2006, citado por SANTOS, CORRÊA, 2009).

O estabelecimento dos Princípios de Paris deu origem à construção de um novo código, o AACR (Anglo-American Cataloging Rules) em 1967, estabelecido pela ALA (American Library Association), pela Canadian Library Association e pela Library

Association da Inglaterra. O AACR foi largamente difundido e adotado em diversos países, e constituiu-se como um grande passo para a uniformização dos catálogos e registros bibliográficos (BARBOSA, 1978).

Entretanto, pode-se destacar que o principal avanço neste período se refere à introdução das tecnologias de informática para a automação dos processos de catalogação e construção dos catálogos.

Durante esse período, por volta dos anos de 1960, as tecnologias de informática começaram a ganhar destaque e ser consideradas importantes na realização de diversos serviços, um deles o processamento da informação. Com isso, em 1965 a LC (Library of Congress) lançou, em fase experimental, o Projeto MARC (Machine Readable Cataloging), considerado na época “[...] uma linguagem padrão para troca de informações bibliográficas”, por isso esse período é denominado de pré-mecanizado (BARBOSA, 1978, p. 46).

De acordo com Mey e Silveira (2009, p. 77-78)

O MARC visava a: 1) aceitação de todos os tipos de materiais; 2) flexibilidade para produção de diferentes aplicativos, além de catálogos; e 3) utilização por diferentes sistemas automatizados.

Existem três pontos a considerar: a) o MARC não é um tipo de catálogo nem um método de catalogação; b) o MARC é um formato, quer dizer, um padrão para entrada e manuseio de informações bibliográficas em computador, não um programa de gerenciamento computacional destas informações; e c) o MARC ajustou os recursos tecnológicos da época à catalogação tradicional, e não o contrário, ou seja, um processo de mecanização (uso da máquina), não ainda de automação. A LC criou também, para o MARC, uma estrutura de organização de dados em fita magnética que se tornou, em 1973, norma internacional, estabelecida pela Organização Internacional de Normalização (ISO) sob o código ISO 2709. [...].

Assim como os códigos de catalogação, o formato MARC passou por inúmeras atualizações ao longo do tempo para contemplar todos os requisitos de descrição de modo automatizado. Não vamos tratar aqui das diferenças entre as versões do MARC, mas é importante destacar que o formato passou a ser amplamente utilizado pela comunidade biblioteconômica por refletir a lógica de descrição contemplada nas estruturas descritivas dos códigos de catalogação, possibilitando, dessa forma, uma facilidade na importação e exportação de dados bibliográficos (FERREIRA, 2002; ALVES, 2005).

Além de serem baseados na estrutura descritiva dos códigos de catalogação, em especial o AACR, os elementos do formato MARC são preenchidos de acordo com as regras externas estabelecidas nos códigos, garantindo, assim, maior consistência na representação por ele gerada (FERREIRA, 2002).

Pode-se considerar que o código de catalogação AACR uniformizou os catálogos e registros bibliográficos, e o formato MARC efetivou a automação desses registros, tornando o processo de catalogação, o processo de intercâmbio e exportação de dados, mais consistente e ágil.

Em relação aos avanços tecnológicos desse período, o que se destaca foi a criação e a difusão dos discos rígidos no final da década de 1960, mudando o cenário do processamento da informação, pois permitiam acesso direto aos dados sem a necessidade de uma sequência ou posição fixa dos dados (SILBERSCHATZ; KORTH; SUDARSHAN, 2006).

[...] Com os discos, podiam ser criados bancos de dados em rede e hierárquicos que permitissem que a estrutura dos dados, como listas e árvores, fossem armazenadas no disco. Os programadores podiam construir e manipular essas estruturas de dados. [...] (SILBERSCHATZ; KORTH; SUDARSHAN, 2006, p. 19).

Dessa forma, os avanços tecnológicos possibilitaram a criação, o desenvolvimento e a implantação do formato MARC, consolidando efetivamente a integração estratégica das tecnologias de informática com o processo de catalogação.

2.5 PERÍODO MECANIZADO

Este período é delimitado pela RIEC até o estabelecimento do Controle Bibliográfico Universal compreendendo o período entre 1969 e meados de 1975 (BARBOSA, 1978). Esta pesquisa considera o período até a década de 1990, devido a importantes acontecimentos que se iniciaram na história da catalogação, principalmente com o surgimento da Internet.

Denomina-se este período como mecanizado, devido à consolidação dos processos de automação da catalogação e dos catálogos em razão do uso efetivo de tecnologias de informática.

A RIEC (Reunião Internacional de Especialistas em Catalogação) foi realizada em 1969 e seu objetivo principal era conseguir, em âmbito internacional, uma maior padronização das regras para a efetivação da catalogação compartilhada com melhor qualidade e a futura disseminação da informação (BARBOSA, 1978).

Os princípios estabelecidos em Paris serviram de base para a mudança de vários códigos; contudo, estavam sendo mal interpretados e até mesmo considerados inconsistentes devido a algumas regras gerais que apresentavam. A catalogação compartilhada da LC

ganhava destaque nessa época e a necessidade de se estabelecer regras internacionalmente aceitas era evidente (BARBOSA, 1978).

Um novo estudo sobre as regras de catalogação foi elaborado por Michael Gorman e o seu trabalho serviu como ponto de partida para a elaboração das ISBDs (International Standard Bibliographic Description – Padrão Internacional para Descrição Bibliográfica) (BARBOSA, 1978). Conforme apontam Mey e Silveira (2009, p. 79):

O notável especialista em catalogação, Michael Gorman, após estudo das informações encontradas em oito bibliografias nacionais correntes, apresentou um documento básico à RIEC, denominado International Standard Bibliographic Description [Descrição bibliográfica internacional normalizada], ou ISBD, que padronizava as informações contidas nas descrições bibliográficas. Para tanto, Gorman sistematizou a ordem das informações e a pontuação utilizada antes de cada informação, de modo a tornar possível seu reconhecimento pelos computadores. A proposta de Gorman, após apreciação internacional, foi publicada em 1971 pela IFLA, como ISBD(M), isto é, para monografias.

Mey e Silveira (2009, p. 79-80) apontam que as ISBDs representaram “[...] um acordo no caminho da padronização. Todos os países se dispuseram a usá-la e essa aceitação internacional acarretou mudanças nos códigos de catalogação, que incorporaram essas normas em novas edições”.

Conforme apontam as autoras, é importante considerar aqui que a padronização não está na quantidade de elementos, mas na forma e na ordem em que a informação é registrada (MEY; SILVEIRA, 2009). Em outras palavras, a padronização se dá pela posição da informação demarcada pela pontuação, pois cada pontuação delimita um tipo de informação a ser descrita, ou seja, um atributo ou característica do recurso. Isso garante a uniformidade dos dados e a padronização necessária para o intercâmbio dos dados de modo automatizado.

Várias resoluções foram determinadas na RIEC, dentre as quais se destaca a proposta de Suzanne Honoré:

[...] esforços deveriam ser conjugados para a criação de um sistema internacional de permuta de informações, pelo qual a descrição bibliográfica normalizada de cada publicação deveria ser estabelecida no seu país de origem e distribuída através de uma agência nacional. Os meios de distribuição deveriam ser fichas ou registros legíveis por máquinas. A eficiência do sistema dependeria da máxima normalização da forma e do conteúdo da descrição bibliográfica. (BARBOSA, 1978, p. 141).

Reforçando esse ponto de vista Doroty Anderson defende o estabelecimento imediato do registro bibliográfico logo após a publicação em seu país de origem, de acordo com normas internacionalmente aceitas para que pudesse ser processado por sistemas manuais ou

mecanizados. Inicia-se a consolidação do CBU - Controle Bibliográfico Universal (BARBOSA, 1978).

É interessante ressaltar que o ideal do CBU não é novo. Conforme aponta Machado (2003), a expressão Controle Bibliográfico (CB) data de 1949 e passou a ser adotada entre os profissionais bibliotecários e documentalistas em 1950. O termo passou a ser utilizado para designar “[...] desde listas de referências até mesmo qualquer atividade ligada à armazenagem e recuperação da informação [...]” (MACHADO, 2003, p. 40).

Os ideais do CBU vão muito ao encontro das idéias estabelecidas por Jewett no período tradicional. Dessa forma, percebe-se que esses ideais foram buscados por diversos profissionais ao longo do tempo, contudo não existiam condições suficientes para que pudessem ser efetivados. Essas condições estavam pautadas basicamente na uniformidade dos registros e no uso de tecnologias de informática.

De acordo com Campello (2006), o conceito de CBU foi formalizado pela IFLA (Federação Internacional de Associações e Instituições Bibliotecárias) em 1974, com a criação do International Office UBC (Escritório Internacional para o Controle Bibliográfico Universal).

[...] Suzanne Honoré [...] definiu as bases de um sistema de intercâmbio de informação que, por meio de agências nacionais, distribuiria os registros bibliográficos padronizados de todas as publicações. A eficiência do sistema, dependeria, portanto, da máxima padronização da forma e do conteúdo da descrição bibliográfica. (CAMPELLO, 2006, p. 02).

No domínio da documentação, várias iniciativas foram elaboradas em relação às bibliografias e à tentativa de consolidação do CBU. O cientista belga Paul Otlet destaca-se na área da documentação como um precursor dos ideais de acesso universal ao conhecimento, contribuindo para o desenvolvimento de várias metodologias de tratamento da informação, mas principalmente vinculando esse tratamento à necessidade de estabelecimento de tecnologias no processo de recuperação dessas informações (ALVES et.al., 2007).

O ideal do CBU é buscado em tempos remotos. Com a invenção da tipografia, foi possível estabelecer com os catálogos impressos um início para a concretização do ideal de CBU. Mas somente com o advento da tecnologia de informática e sua consolidação nos processos de tratamento descritivo da informação foi possível instituir as bases para a efetivação do CBU na prática. Dessa forma, a produção bibliográfica mundial passou a ser acessada pelas redes de informação, em meados nos anos de 1970; iniciou-se o processo de conversão retrospectiva dos registros de bibliográficos, em meados dos anos de 1980; e a

disponibilização universal dos catálogos das bibliotecas pela Internet, em meados dos anos de 1990 (CAMPELLO, 2006).

Conforme se percebe, o acesso mundial aos registros bibliográficos foi possível não só pelo estabelecimento do CBU, como também por meio de condições que o efetivassem, tais como a consolidação de um tratamento descritivo da informação, que garantisse a uniformidade dos registros, aliado ao uso de tecnologias de informática.

Neste período as teorias, os princípios e os instrumentos para o tratamento descritivo da informação já foram estabelecidos e o formato MARC se estabeleceu como formato padrão para representação informacional no domínio bibliográfico.

Destaca-se o desenvolvimento de grandes bancos de dados bibliográficos e redes cooperativas, assim ressalta-se tecnologicamente o desenvolvimento dos primeiros sistemas de bancos de dados relacionais comerciais, fato que desencadeou grande avanço no desenvolvimento de técnicas de processamento e consultas declarativas. Além disso, este período foi marcado pelo desenvolvimento e aprimoramento dos modelos para construção de bancos de dados (modelo relacional, hierárquico, de rede, entre outros) e pelo desenvolvimento de linguagens para a construção de bancos de dados (SILBERSCHATZ; KORTH; SUDARSHAN, 2006).

Sendo assim, pode-se considerar que este período foi caracterizado nitidamente pela integração estratégica das tecnologias de informática com os processos de TDI.

Entretanto, a partir da década de 1990, o domínio bibliográfico passou por intensas mudanças causadas pelos avanços das tecnologias de informática, em especial com o desenvolvimento das TICs e o surgimento da Internet.

Essas mudanças colocaram em discussão a estrutura vigente de disponibilização e acesso às informações, devido ao surgimento crescente de novos tipos de recursos informacionais; novas formas para o tratamento descritivo da informação começaram a ser estudadas e desenvolvidas.

2.6 PERÍODO DE METADADOS

Este período inicia-se nos anos de 1990 e se estende até os dias atuais. É marcado principalmente pelo uso intenso de tecnologias de informática, que se fazem presentes desde a

produção da informação e seu armazenamento, passando pelos processos de TDI, até chegarem na busca e na recuperação da informação com os metadados.

Caracteriza-se pelo uso intensivo de tecnologias de informática, tanto em relação a hardwares como em softwares, causando impacto nos processos de produção, armazenamento, disponibilização, acesso, localização, busca e recuperação das informações em meio digital.

À medida que as tecnologias de informática e em especial as TICs foram sendo desenvolvidas e aperfeiçoadas, surgiram novos tipos de ambientes informacionais em meio digital e com eles uma crescente variedade de recursos informacionais disponibilizados, fazendo com que os sistemas de recuperação da informação desenvolvessem maneiras de melhorar a busca e a recuperação dessas informações.

Assim, presencia-se a partir deste período a procura por soluções aos problemas de busca e recuperação da informação em ambientes digitais. Essas soluções estão pautadas nos métodos tradicionais de tratamento descritivo da informação e principalmente no desenvolvimento de novas ferramentas tecnológicas, ou uma infraestrutura tecnológica, como: linguagens de marcação, ferramentas para a construção de bancos de dados, entre outros.

Essa “nova” forma de representação informacional, que passou a ser denominada metadados, gerou uma variedade de padrões para representação das informações em meio digital e atualmente os métodos de organização, tratamento e representação informacional estão vinculados ao seu desenvolvimento e aplicação.

É importante lembrar, conforme aponta Méndez Rodríguez (2002), que os problemas de acesso e principalmente de recuperação da informação não são recentes nem estão exclusivamente vinculados ao ambiente informacional digital, conforme pode ser comprovado pelo breve histórico dos catálogos e da catalogação apontado anteriormente.

Dessa forma, a recuperação da informação vem sendo objeto de investigação principalmente nas áreas de Ciência da Informação e Ciência da Computação, que convergem na busca de soluções, criando uma infraestrutura tecnológica e representacional para diversos ambientes científicos: a Ciência da Informação desenvolvendo métodos e técnicas de organização, tratamento e representação informacional; e a Ciência da Computação proporcionando o desenvolvimento de ferramentas tecnológicas para a construção de ambientes informacionais e a construção de sistemas de recuperação da informação mais eficientes (ALVES et al., 2007; ALVES, 2005).

Conforme se observa nos subcapítulos anteriores, os catálogos constituem-se como instrumentos de recuperação da informação nos quais convergem os aspectos

representacionais e tecnológicos presentes em cada época. Dos catálogos manuais às bases de dados bibliográficos em meio digital (OPACs – Online Public Access Catalogues), a história dos catálogos é permeada pelo desenvolvimento dos processos de catalogação, em especial da sistematização de regras de representação da informação, seu aprimoramento e aplicação da mesma tendo como base o uso das tecnologias vigentes em cada época para a otimização de suas tarefas. Assim, ao longo do tempo a Catalogação vem fazendo especial uso das tecnologias de informática e das TICs, proporcionando nas últimas décadas o surgimento de programas de cooperação e compartilhamento de registros bibliográficos, baseados no intercâmbio de dados bibliográficos e catalográficos, garantindo ao usuário o acesso a um número maior de informações.

O período é marcado pelo uso intensivo de tecnologias no processamento e disponibilização das informações. Os sistemas de bancos de dados passaram a oferecer acesso constante a grandes quantidades de informações pela Web e isso exigiu um aprimoramento das tecnologias envolvidas no processamento da informação proporcionando a criação de bancos de dados mais estruturados. (SILBERSCHATZ; KORTH; SUDARSHAN, 2006).

Contudo, o constante aprimoramento das tecnologias de informática, em especial das TICs, continua a redefinir as tradicionais atividades de produção, armazenamento, tratamento, busca, localização, recuperação, acesso e uso das informações, fazendo com que os profissionais envolvidos com o processamento da informação desenvolvam e aprimorem métodos de tratamento descritivo da informação mais condizentes com os novos ambientes digitais.

Conforme aponta Méndez Rodríguez (2002), os metadados protagonizam um novo paradigma nos sistemas de informação digital deste milênio. Contudo, mesmo sendo criados para o contexto tecnológico, muitos dos padrões de metadados desenvolvidos não atendem satisfatoriamente às necessidades de representação informacional em um domínio específico, pois seus esquemas de descrição são amplos e gerais e por isso não apresentam especificidade do domínio. Deste modo, presencia-se atualmente o delinear de um novo panorama, pautado no uso intensivo de tecnologias de informática e no estabelecimento de princípios mais consistentes para a determinação de metadados para que representem um recurso informacional com unicidade, tanto no domínio da Web como em domínios específicos, como o domínio bibliográfico.

Neste sentido, antes de tratarmos das tendências atuais sobre o TDI e sabendo que essa tendência está vinculada aos metadados, é necessária uma revisão dos aspectos principais dos metadados, já que se constituem como elementos fundamentais para proporcionar o

tratamento descritivo informacional no contexto digital. Esses aspectos serão tratados no capítulo 3 a seguir.

2.7 SISTEMATIZAÇÃO DA INTEGRAÇÃO ESTRATÉGICA DA TDI COM AS TECNOLOGIAS

Com base nos estudos de Barbosa (1978), Machado (2003), Campello (2006), Chan (2007), Mey e Silveira (2009), Santos e Corrêa (2009), Silberschatz, Korth e Sudarshan (2006), apresenta-se a seguir um quadro sistematizando os principais acontecimentos relacionados à consolidação do tratamento descritivo da informação e sua integração estratégica com as tecnologias. Optou-se por indicar apenas os períodos que mais trouxeram contribuições significativas em relação ao TDI e aos avanços tecnológicos.

PERÍODOS HISTÓRICOS	CONTRIBUIÇÕES PARA A TEORIA DA CATALOGAÇÃO	DESENVOLVIMENTO DE PADRÕES	INTEGRAÇÃO ESTRATÉGICA COM AS TECNOLOGIAS
Períodos Remotos	Catálogo de Saint Martin: dividido em três seções, a última das quais incluía análise das partes (início das entradas analíticas)	Não havia uma forma de organização e sistematização unificada das informações	Suportes informacionais: tabletes de argila, madeira, papiro, pergaminho, papel.
Período Tipográfico	Primeiras iniciativas para construção de um código de catalogação – sem padronização unificada e sem consolidação de princípios	Primeiros catálogos livreiros – sem padronização unificada	Tipografia: catálogos em forma de livros impressos; catálogos em ficha
Período Tradicional	Princípios de Panizzi Princípios de Jewett Princípios de Cutter Princípios de Lubetsky	Regras de Catalogação do Museu Britânico (1839) Regras de Jewett (1853) Regras de Cutter (1876) AA (1908) Instruções Prussianas Código da Vaticana ALA (1941 Draft) LC (1949) ALA (1949)	Tecnologias mecânicas: início dos catálogos automatizados (sistema de processamento Hollerith de cartões perfurados); tecnologias de armazenamento – fitas magnéticas
Período Pré-Mecanizado	Princípios de Paris (1961)	AACR (1967) MARC (1965)	Tecnologias de informática: construção de bancos de dados para processamento da informação; tecnologias de armazenamento – disco rígido
Período Mecanizado	RIEC (1969) CBU (Controle Bibliográfico Universal)	ISBD (International Standard Bibliographic Description – Padrão Internacional para Descrição Bibliográfica) (1971) AACR2 (1978) AACR2R (1988)	Tecnologias de informática: consolidação da automação bibliográfica, consolidação de grandes bancos de dados, redes cooperativas, OPACs; desenvolvimento e aprimoramento dos modelos e linguagens de construção de bancos de dados
Período de Metadados	Desenvolvimento dos Padrões de metadados	Padrão de metadados Dublin Core (1995) AACR2R (1998) AACR2R (2002) MARC 21 e desdobramentos	Tecnologia de informática e TICs: sistemas de recuperação da informação, bancos de dados com acesso à Web, aprimoramento das tecnologias para a construção dos bancos de dados

QUADRO 1: Sistematização da história da Catalogação em períodos.

FONTE: do autor.

No capítulo 6 serão retomadas algumas questões sobre o desenvolvimento do TDI e sua integração com as tecnologias, com o intuito de atingir os objetivos propostos de categorização dos instrumentos para o TDI no domínio bibliográfico.

3 METADADOS: revisando fundamentos, princípios, conceitos, tipologia, funções, características e aplicações

Conforme apontado anteriormente, os metadados estão presentes nos sistemas de informação em uma variedade de tipos e níveis. Sua construção adequada e aplicação do padrão de metadados correspondente ao tipo de ambiente informacional garantem maior efetividade nos sistemas e, conseqüentemente, uma recuperação com melhores resultados.

Contudo, os metadados vêm sendo alvo de intensos estudos e debates científicos por todas as peculiaridades abarcadas pelo termo, não só no nível conceitual (definições, características, tipologia e função), mas também no nível de desenvolvimento (criação, normalização, aplicação e uso) e, principalmente, pelos aspectos de mudanças que ocasionam.

Embora existam vários estudos sobre os metadados, é necessário desenvolver uma revisão sobre alguns de seus aspectos e uma análise crítica de alguns pontos de vista, pois pretende-se verificar a relação do tratamento descritivo da informação da área de Ciência da Informação com os metadados.

Dessa forma, busca-se aqui elucidar: como podemos definir atualmente os metadados, tendo em vista seu aspecto tecnológico e representacional? Qual sua abrangência e seu escopo? Quais suas características e aplicações? Como vem sendo seu desenvolvimento e padronização? Qual sua relação com a Web e com a Ciência da Informação? Por que são considerados importantes para os sistemas de informação? Qual é a mudança de paradigma que proporcionam? Buscou-se responder a essas questões no decorrer deste capítulo, com o intuito de se chegar a uma contextualização teórica para solucionar o problema de pesquisa.

Tendo como base os estudos anteriormente desenvolvidos sobre os metadados e a literatura levantada para a realização da pesquisa, procurou-se destacar aqui os aspectos considerados mais relevantes.

Um breve panorama histórico do termo metadados demonstra que a concepção do termo antecede a Web, tendo sido criado por Jack E. Myers nos anos de 1960 para descrever “conjuntos de dados”. Embora criado inicialmente como uma marca comercial, o termo metadados começou a se desenvolver como teoria para designar a descrição de recursos informacionais, passando a ser utilizado amplamente em diversas áreas do conhecimento. Na década de 1980 o termo metadados começa a surgir na literatura de sistemas de gestão de base de dados (SGBD) para designar os dados que documentavam as características da informação

contida nas bases de dados. Na década de 1990 o termo metadados começou a ser utilizado e aplicado em comunidades específicas, destacando-se a área de dados geoespaciais, na qual se pode considerar uma das primeiras aplicações dos metadados e desenvolvimento de padrões de metadados para gestão e interoperabilidade de dados geoespaciais em meio digital (CSDGM – Content Standards for Digital Geospatial Metadata, desenvolvido pela FGDC – Federal Geographic Data Committee). Em 1995, com o desenvolvimento do padrão de metadados Dublin Core, o termo metadados começou a ser utilizado para designar a descrição de recursos informacionais na Web (MÉNDEZ RODRÍGUEZ, 2002, SENSO; ROSA PIÑERO, 2003; ZENG, QIN, 2008).

Entretanto, é importante destacar outra vertente encontrada na literatura sobre a concepção dos metadados, ressaltando que sua existência independe do ambiente informacional digital. Conforme aponta Ortiz-Repiso Jiménez (1999), os metadados em si não são algo novo, a novidade está na variedade de padrões disponíveis e em desenvolvimento e também nas formas como vêm sendo utilizados. Nesse sentido, os autores Milstead e Feldman (1999) apontam que desde as primeiras tentativas de organização da informação na história da Biblioteconomia, catalogadores e indexadores criam e padronizam metadados em suas atividades de tratamento e representação informacional: “[...] padrões de informação bibliográfica, sumários, termos de indexação, e abstracts são todos substitutos do material original, portanto metadados”.

Gilliland-Swetland (1999) aponta que desde os anos de 1960 as bibliotecas vêm compartilhando metadados descritivos em sistemas informatizados, por meio do uso de normas e estruturas de descrição, tal como o formato MARC.

Apesar de o termo metadados ter sido amplamente difundido para designar representação informacional em meio digital, essa vertente demonstra que a concepção para o termo metadados independe do ambiente digital. Dessa forma, pode-se afirmar que ao longo do tempo a Biblioteconomia vem criando metadados por meio do desenvolvimento das teorias, dos princípios, dos fundamentos e dos métodos de Tratamento Descritivo da Informação sem, contudo, utilizar essa denominação e antes da utilização das tecnologias de informática.

O termo metadados significa literalmente “dados sobre dados”, conforme mencionam Miller (1996); Souza, Catarino e Santos (1997); Milstead e Feldman (1999); Gilliland-Swetland (1999); Souza, Vendrusculo e Melo (2000); Takahashi (2000); Senso e Rosa Piñero (2002); Méndez Rodríguez (2002); Zeng e Qin (2008) entre outros.

Contudo, não há um consenso sobre um conceito único para o termo. Ao mesmo tempo em que se apresenta como um termo neutro às diversas áreas do conhecimento que trabalham com tratamento e processamento informacional (SENSO; ROSA PIÑERO, 2003), o uso do termo evidencia uma ambiguidade designativa e uma pluralidade de definições que variam de acordo com o contexto de uso. Dessa forma, procurou-se aqui revisar o significado do termo e suas acepções para a área de Ciência da Informação (GILLILAND-SWETLAND, 1999; MÉNDEZ RODRÍGUEZ, 2002; ROSETTO, 2003).

Assim, para melhor entendimento do termo metadados, buscou-se aqui abordar algumas definições e considerações sobre o termo no intuito de refletir sobre seu significado e propor uma nova definição, tendo em vista os aspectos informacionais da Ciência da Informação e tecnológicos da Ciência da Computação.

De acordo com Méndez Rodríguez (2002), os metadados vêm sendo utilizados como termo descritivo por si só, mas apresentam um significado reduzido e um pouco vago quando não se contextualiza seu domínio, conteúdo e âmbito de aplicação. Para a autora, a definição “dados sobre dados” apresenta-se de forma generalizada e pode ser válida em qualquer contexto; ressalta ainda que os metadados podem ser considerados apenas como dados quando são utilizados para falar deles mesmos.

Outro ponto de vista destacado por Méndez Rodríguez (2002) refere-se à análise da palavra metadados. “Meta” é um “prefixo que indica que algo se aplica a si mesmo [...]”; por esse motivo pode-se considerar os metadados “[...] também dados, mas dados representacionais, que acrescentado a própria informação adquirem um valor semântico para substituí-la ou representá-la” (BERNERS-LEE, 2000, p. 225, citado por MÉNDEZ RODRÍGUEZ, 2002, p. 30).

Semelhante a esse ponto de vista, encontram-se as seguintes definições no âmbito de aplicação de sistemas de base de dados computacionais: “Metadados são *dados que descrevem outros dados*” (TAKAHASHI, 2000, p. 59), ou ainda “Metadados são descrições de dados armazenados em um banco de dados [...]” (SOUZA; CATARINO; SANTOS, 1997, p.02)

Em um sistema de banco de dados os metadados são atributos², ou seja, são as características que identificam e representam uma entidade (objeto do mundo real que pode ser, por exemplo, um recurso informacional); em outras palavras, qualquer propriedade,

² “Um atributo é um dado (informação de estado) para qual cada Objeto em uma Classe tem seu próprio valor” (COAD, YOURDON, 1991, p. 117). Essa definição provém da área de Análise Orientada a Objeto e se constitui com conceitos adotados em sistemas de informações.

qualidade ou característica atribuída para descrever um objeto do mundo real (dados representacionais) em um sistema de banco de dados (CHEN, 1990; COAD, YOURDON, 1991).

Encontram-se na literatura vários autores que explicam mais detalhadamente o termo metadados. Serão apresentadas a seguir algumas definições consideradas mais importantes para a área de Ciência da Informação:

Metadados, para Ortiz-Repiso Jiménez (1999, p. 218), podem ser definidos como

[...] um conjunto de dados que pode ser usado para descrever e representar recursos informacionais. Contém um conjunto de elementos de dados que podem ser usados para descrever o conteúdo e a localização de um recurso informacional e facilitar sua recuperação e acesso na rede.

Para Takahashi (2000, p. 172), metadados são

Dados a respeito de outros dados, ou seja, qualquer dado usado para auxiliar na identificação, descrição e localização de informações. Trata-se em outras palavras, de dados estruturados que descrevem as características de um recurso de informação.

Os autores Souza, Vendrúsculo e Melo (2000, p. 93) definem metadados como “[...] dado sobre dado. É a catalogação do dado ou descrição do recurso eletrônico”.

Senso e Rosa Piñero (2002, p. 99) consideram os metadados como

[...] toda aquela informação descritiva sobre o contexto, qualidade, condição ou características de um recurso, dado ou objeto que tem a finalidade de facilitar sua recuperação, autentificação, evolução, preservação ou interoperabilidade.

Grácio (2002, p. 21) afirma:

Comumente chamado de dados sobre dados, o termo metadados pode ser melhor descrito como um conjunto de dados chamados de elementos, cujo número é variável de acordo com o padrão, e que descreve o conteúdo de um recurso, possibilitando a um usuário ou a um mecanismo de busca acessar e recuperar esse recurso. Esses elementos descrevem informações do tipo nome, descrição, localização, formato, entre outras, que possibilitam um número maior de campos para pesquisas.

Em outras palavras, metadados são um “conjunto de elementos que descrevem as informações contidas em um recurso, com o objetivo de possibilitar sua busca e recuperação” (GRÁCIO, 2002, p. 23).

Analisando as definições encontradas na literatura, observa-se que os metadados estão diretamente vinculados a estruturas padronizadas de descrição e aos objetivos que se pretende representar com sua aplicação em um sistema de um determinado domínio. De modo geral, pode-se considerar os metadados dados estruturados e padronizados que descrevem um

recurso informacional, com o objetivo de facilitar sua identificação para localização, busca e recuperação em um sistema ou ambiente informacional (ALVES, 2005).

Conforme verificado em algumas definições, o termo metadados está intimamente associado com a definição de padrões de metadados. Para que eles possam existir, os metadados (*metadata*) devem estar codificados em estruturas padronizadas de descrição, denominadas como padrões de metadados (*metadata statement*). O conjunto de metadados ou elementos de metadados (*element sets*) irá compor o esquema de metadados (*metadata schema*) do padrão de metadados.

A autora Rosetto (2003, p. 59) destaca a diferença entre metadados e padrões de metadados, estes últimos denominados também formatos de metadados:

Metadados são um conjunto de dados – atributos – referenciais, metodologicamente estruturados e codificados, conforme padrões internacionais, para localizar, identificar e recuperar pontos informacionais de textos, documentos e imagens disponíveis em meios digitais ou em outros meios convencionais, e

Formatos de metadados referem-se a padrões que estabelecem regras para a definição de atributos (metadados) de recursos de informacionais, para a) obter coerência interna entre os elementos por meio de semântica e sintaxe; b) promover necessária facilidade para esses recursos serem recuperados pelos usuários; c) permitir a interoperabilidade dos recursos de informação.

De modo resumido, pode-se também considerar os metadados e padrões de metadados de acordo com os estudos de Alves (2005), que considera os metadados atributos ou dados referenciais que representam um recurso, e o conjunto codificado e padronizado de metadados (atributos) constitui o padrão de metadados ou formato de metadados.

Analisando as definições encontradas na literatura e tendo como base principal os autores Chen (1990) e Alves (2005), foi possível elaborar uma sistematização do conceito para os termos metadados e padrões de metadados com o intuito de atualizar e redefinir o significado desses termos para esta pesquisa, conforme apontado a seguir:

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.

É importante destacar que um metadado ou atributo irá representar uma única característica da entidade, enquanto que o conjunto de metadados ou o conjunto de atributos presentes em um padrão irá representar a entidade no todo. É o conjunto de metadados que irá identificar e individualizar a entidade de forma única e inequívoca³ entre as demais. Dessa forma, os metadados ou os atributos escolhidos para representar uma entidade devem ser capazes de identificá-la de forma absoluta, ou seja, garantindo a sua unicidade (CHEN, 1990).

Diante do exposto, pode-se considerar que para o domínio bibliográfico os metadados e padrões de metadados estão relacionados com o tratamento descritivo da informação (TDI), pois garantem a descrição e a representação padronizada do recurso informacional independente do tipo de suporte, com o objetivo de facilitar a busca e a recuperação, aspecto que vem sendo estudado há muito tempo na Biblioteconomia e também na Ciência da Informação (GILLILAND-SWETLAND, 1999; ALVES, 2005).

3.1 TIPOS, CARACTERÍSTICAS, FUNÇÕES, APLICAÇÕES E IMPORTÂNCIA DOS METADADOS

Há uma diversidade de tipos de metadados que variam de acordo com o recurso informacional a ser representado, com domínio de aplicação e com as necessidades dos usuários. Assim como se constatou a diversidade de definições para o termo metadados, é possível encontrar na literatura várias categorizações sobre os tipos de metadados existentes. Desta forma, serão apresentados aqui alguns tipos considerados mais relevantes para esta pesquisa.

Os autores Gilliland-Swetland (1999), Senso e Rosa Piñero (2003) e Rosetto (2003) destacam que os tipos de metadados estão relacionados com as características e funções que eles representam. Deste modo, apresenta-se a seguir a classificação de alguns tipos de metadados conforme as funções que desempenham:

- Administrativos: são metadados usados no gerenciamento e administração dos recursos de informação. Esse tipo de metadado fornece informações como: data de criação dos

³ Em que não há equívoco; claro, evidente, manifesto. (FERREIRA, 2004).

recursos, tipos de arquivos, formas de acesso, controle de direitos e reproduções, informação sobre registros legais, informação sobre localização etc (GILLILAND-SWETLAND, 1999; ROSETTO, 2003; SENSO, ROSA PIÑERO, 2003);

- Descritivos: são metadados usados para descrever, identificar e representar recursos de informações. Esse tipo fornece informações relacionadas com a catalogação, como título, autor, imprensa, data, resumo, palavras-chave, e ainda a relação dos hiperlinks entre os recursos, anotações de usuários etc (GILLILAND-SWETLAND, 1999; ROSETTO, 2003; SENSO, ROSA PIÑERO, 2003);
- Conservação: são metadados relacionados com a conservação e preservação dos recursos de informação. Esse tipo fornece informações sobre as condições físicas de um recurso, informações de como conservar e preservar as versões físicas e digitais de um recurso etc (GILLILAND-SWETLAND, 1999; ROSETTO, 2003; SENSO, ROSA PIÑERO, 2003);
- Técnicos: são metadados relacionados com o funcionamento dos sistemas e o comportamento dos metadados. Esse tipo fornece informações sobre hardware e software, digitalização, controle do tempo de resposta dos sistemas, autenticidade e segurança dos dados etc (GILLILAND-SWETLAND, 1999; ROSETTO, 2003; SENSO, ROSA PIÑERO, 2003);
- Uso: são metadados relacionados com o nível e tipo de uso dos recursos de informação. Esse tipo fornece informações sobre os registros de exibição, controle de uso e usuários, controles de acesso, informação sobre versões múltiplas etc (GILLILAND-SWETLAND, 1999; ROSETTO, 2003; SENSO, ROSA PIÑERO, 2003).

Esses tipos de metadados apresentados se constituem como uma tipologia geral; dentro de cada um existem metadados específicos para a representação das características de uma entidade, codificados de acordo com um padrão de metadados. Dessa forma, um padrão de metadados pode conter tipos de metadados em mais de uma dessas categorias.

Os tipos de metadados apresentados estão separados em categorias de funcionalidade, mas, além disso, existem outros fatores que caracterizam os metadados derivados do complexo processo de criação manual e automática e por outras diferentes funções. Os autores Gilliland-Swetland (1999) e Senso e Rosa Piñero (2003) destacam os seguintes fatores:

- Fonte dos metadados: se os metadados são internos, gerados no momento da criação do recurso (exemplo: nomes de arquivos), ou são metadados externos, gerados posteriormente à criação do recurso (exemplo: fichas e registros de catalogação);

- Método para criação dos metadados: se os metadados são automáticos, gerados automaticamente por um computador (exemplo: índices de palavras-chave), ou são metadados manuais, criados por indivíduos (exemplo: registros de catalogação);
- Caráter dos metadados: se os metadados foram criados por indivíduos que não são especialistas da área de informação (exemplo: metadados criados pelo desenvolvedor de uma página pessoal), ou são metadados criados por especialistas temáticos ou da área de informação (exemplo: registros em formato MARC elaborados por um bibliotecário);
- Status: se os metadados são estáticos, que não mudam depois de criados (exemplo: título e data de criação de um recurso), ou são metadados dinâmicos, que podem se modificar de acordo com o uso e a manipulação do recurso (exemplo: registros de operações dos usuários); se os metadados são de longa duração, para assegurar o acesso e o uso do recurso (exemplo: formatos técnicos e processamento da informação), ou são metadados de curta duração, que são principalmente do tipo operacional (exemplo: informam sobre conservação e administração dos recursos);
- Estrutura: se os metadados são estruturados, ou seja, apresentam uma estrutura previsível, pré-determinada, baseada em um padrão normalizado internacionalmente (exemplo: MARC), ou são metadados não estruturados, ou seja, não possuem estrutura previsível (exemplo: formatos de banco de dados locais);
- Semântica: se os metadados são controlados, ou seja, aqueles que seguem ou são normalizados por um vocabulário controlado, formulário de autoridade etc (exemplo: AACR2), ou são metadados não controlados, ou seja, aqueles que não seguem um vocabulário controlado ou formulário de autoridade (exemplo: meta etiquetas HTML);
- Nível: se os metadados são de coleções, ou seja, estão relacionados a coleções de documentos ou recursos (exemplo: uma coleção de recursos descritos pelo formato MARC), ou são metadados individuais, ou seja, metadados relacionados com recursos individuais ou que não pertencem a nenhuma coleção (exemplo: legenda de uma imagem).

Conforme observado, os metadados existem sob diferentes tipos para atender a diferentes funções em um domínio de aplicação. As categorias tipológicas apresentadas foram consideradas mais relevantes para esta pesquisa; contudo, é importante deixar claro que de modo geral todos os tipos de metadados se constituem como pontos de acesso aos recursos informacionais que representam, sejam os metadados intrínsecos às entidades (aqueles atributos que individualizam a entidade), sejam extrínsecos (aqueles atributos que identificam a entidade em um sistema) (MÉNDEZ RODRÍGUEZ, 2002).

Os autores Zeng e Qin (2008, p. 08) consideram que atualmente os metadados abarcam os seguintes benefícios: os usuários podem navegar do registro de metadados até o recurso na Internet, os registros de metadados indicam a localização física do recurso informacional, e ainda há a possibilidade de processamento automático de certos tipos de metadados minimizando os custos com o processamento manual.

Gilliland-Swetland (1999) aponta que os metadados apresentam outras vantagens: auxiliam na contextualização do recurso no sistema; mantêm as relações entre recursos informacionais; possibilitam a distinção entre versões múltiplas de um mesmo recurso; permitem representar questões legais, privacidade, interesse de propriedade, restrições de uso, direitos autorais etc; possibilitam conservar informações que caracterizam os recursos em atualizações de sistemas, entre outras vantagens. Assim, considera-se que os metadados não apenas descrevem uma entidade como também a contextualizam em um sistema de informação, garantindo seu gerenciamento.

Quanto à função dos metadados, pode-se dizer que ela está relacionada com o propósito que se pretende atingir com sua aplicação em um sistema de informação.

De modo geral, a primeira função é garantir o gerenciamento da informação em um sistema, por meio da identificação, descrição e controle da informação; para isso, desempenha alguns papéis de acordo com o propósito de aplicação. Conforme aponta Méndez Rodríguez (2002, p. 48), observam-se dois papéis ou perspectivas em relação à aplicação de metadados: uma tradicional e uma emergente, conforme explicado a seguir:

PAPÉIS TRADICIONAIS	PAPÉIS EMERGENTES
1. Identificação e descrição da informação	4. Autoria e propriedade intelectual
2. Busca e recuperação	5. Formas de acesso
3. Localização dos documentos	6. Atualização da informação
	7. Preservação e conservação
	8. Restrição de uso
	9. Valoração do conteúdo
	10. Visibilidade da informação
	11. Acessibilidade dos conteúdos

QUADRO 2: Dupla perspectiva do papel dos metadados.
FONTE: Méndez Rodríguez (2002, p. 48).

Essas aplicações demonstram a crescente importância dos metadados para os sistemas de informação, principalmente em relação ao gerenciamento de informação digital, lembrando que nenhuma dessas aplicações exclui a outra. Dessa forma, a função dos metadados em um sistema de informação digital ocorre das seguintes formas: (a) exerce a mesma função dos catálogos e sistemas de descrição bibliográfica, com identificação, descrição, busca, localização e recuperação da informação; (b) em relação aos sistemas, sua função é facilitar a interoperabilidade e intercâmbio de dados, facilitar a busca e o gerenciamento da informação; (c) em relação aos usuários, indicar a informação acessível, sua localização e as condições para se chegar até ela (MÉNDEZ RODRÍGUEZ, 2002).

Como pode ser observado, existe uma variedade de funções para os metadados; contudo, para esta pesquisa, a principal função dos metadados está relacionada com a otimização da recuperação da informação. A aplicação apropriada dos metadados, ou seja, sua construção padronizada e seu uso preciso proporcionam uma melhora na recuperação da informação nos sistemas de informação digital, aumentando a probabilidade de os usuários encontrarem informações relevantes para suas necessidades.

Dessa forma, em relação à Ciência da Informação, pode-se dizer que a principal função dos metadados é fornecer uma representação padronizada e inequívoca dos recursos informacionais, multiplicando as formas de acesso ao recurso, com a finalidade de busca, localização, recuperação, intercâmbio de dados e interoperabilidade entre sistemas.

Os metadados são considerados importantes, principalmente pelo fato de proporcionarem uma economia e melhora do funcionamento dos sistemas. Entretanto, é preciso deixar claro que eles são inerentes aos sistemas de informação, seu uso é não somente importante como fundamental para a existência dos sistemas.

Assim, pode-se considerar que a importância dos metadados para os sistemas de informação consiste na aplicação coerente dos padrões de metadados e outros instrumentos como os esquemas de codificação, que garantam a construção padronizada de representações.

3.2 PRINCÍPIOS GERAIS

Conforme apontado anteriormente, a função dos metadados está relacionada com o propósito que se pretende atingir com sua aplicação em um sistema de informação.

Entretanto, para serem úteis em um sistema, é preciso que seu desenvolvimento esteja pautado em princípios de modo que sua criação não seja feita de modo arbitrário.

A aplicação ampla de metadados em diferentes domínios ou comunidades de interesse gerou diferentes práticas e diferentes estruturas de padrões de metadados baseadas em necessidades próprias e regras próprias de representação (ZENG, QIN, 2008).

Porém, para que o desenvolvimento de padrões de metadados pudesse ocorrer de forma unificada, foram sendo desenvolvidos alguns princípios que nortearam sua criação, produção e estabelecimento e, conseqüentemente atuam sobre as características da estrutura de descrição (conjunto de metadados descritivos do padrão).

Embora cada padrão de metadados seja criado a partir de um princípio específico do domínio ao qual ele pertence (abordado no capítulo 4), é possível determinar princípios gerais ou diretrizes para o estabelecimento dos metadados e dos padrões de metadados.

Neste tópico serão tratados alguns princípios gerais para a construção de metadados, e como foco desta pesquisa serão tratados mais especificamente os metadados descritivos.

A NISO - *National Information Standards Organization* -, organização sem fins lucrativos credenciada pela ANSI - *American National Standards Institute* -, responsável por identificar, desenvolver, manter e publicar normas técnicas para o gerenciamento da informação em meio digital, estabelece algumas diretrizes ou princípios para a construção de metadados (NISO, 2004).

Dentre as normas criadas apresenta um guia para a construção de coleções digitais (*“Framework of Guidance for Building Good Digital Collections”*), no qual apresenta seis princípios para a aplicação de metadados, destacando os requisitos desejáveis para a construção de metadados. Originalmente é utilizado o termo “good metadata” que aqui se entende como metadados de qualidade (NISO, 2004, p. 10, tradução nossa):

- Metadados de qualidade devem ser apropriados aos materiais de uma coleção, aos usuários da coleção e destinados ao provável uso por objetos digitais;
- Metadados de qualidade devem ser apropriados para proporcionar a interoperabilidade;
- Metadados de qualidade devem proporcionar a utilização de padrões como vocabulários controlados, para refletir o que, onde, quando e quem em relação ao conteúdo;
- Metadados de qualidade devem incluir declarações claras sobre as condições e termos de uso do recurso informacional;
- Registros de metadados de qualidade são objetos em si e, portanto, devem ter a qualidade de arquivamento, persistência e identificação única. Metadados de qualidade devem ser confiáveis e verificáveis;
- Metadados de qualidade apóiam a gestão a longo prazo de recursos em coleções.

Além disso, há várias iniciativas em andamento para estabelecer melhor qualidade aos metadados, consistindo no estabelecimento de princípios gerais para a construção de metadados, tais como:

- Aperfeiçoamento dos *templates* (planilhas) dos padrões de metadados, apresentando seleção de campos e melhoramento nas regras de validação;
- Desenvolvimento e aprimoramento dos softwares de interoperabilidade e *crosswalk* entre diferentes padrões;
- Criadores de conteúdo são capacitados para entender os metadados, conceitos de vocabulários controlados e uso de ferramentas de metadados;
- Vocabulários controlados criados para um domínio específico estão ficando mais conhecidos e com uso mais amplo;
- Comunidades de usuários estão desenvolvendo e aperfeiçoando esquemas de metadados, perfis de aplicação, vocabulários controlados e diretrizes de uso.

Em 2001, o *Dublin Core Metadata Initiative (DCMI)* e o *Institute for Electrical and Electronics Engineers (IEEE) Learning Object Metadata (LOM) Working Group* se reuniram para estabelecer alguns princípios considerados básicos e comuns a todos os domínios de metadados, auxiliando no estabelecimento de padrões de metadados e aplicações.

Atualmente o DCMI é a instituição responsável por guiar essa construção de metadados e padrões de metadados, estabelecendo princípios comuns que devem estar em consonância com regras e normas emergentes e internacionais, com o intuito de garantir a construção adequada e padronizada de metadados e, com isso, a interoperabilidade entre sistemas (ZENG, QIN, 2008).

Tendo como base as resoluções do DCMI, os autores Zeng e Qin (2008) destacam a existência de três princípios básicos para a construção de metadados e que se constituem como requisitos de um sistema de informação: simplicidade, extensibilidade e interoperabilidade. Cada um desses princípios será explicado a seguir.

- *Simplicidade*: refere-se aos atributos (metadados) necessários para a manutenção de um conjunto mínimo de elementos descritivos para facilitar a implementação. É importante que os metadados sejam flexíveis para acomodar as necessidades específicas de descrição, permitindo que sejam incluídos novos metadados para atender a necessidades específicas de um domínio específico (ZENG; QIN, 2008, p. 10, tradução nossa).
- *Extensibilidade*: pode ser entendida de duas formas; a capacidade de um padrão de metadados oferecer um conjunto de elementos descritivos que possa unificar os diferentes

padrões de descrição; e a capacidade de intercâmbio entre registros de metadados de um padrão de metadados mais simples para outro mais complexo. A extensibilidade requer que os sistemas permitam a adição de novos elementos de metadados em um padrão baseados em normas existentes, ou elementos estabelecidos em nível local (ZENG; QIN, 2008, p. 10, tradução nossa).

- *Interoperabilidade*: “é a capacidade de múltiplos sistemas com diferentes hardwares e plataformas de softwares, estruturas de dados e interfaces intercambiarem dados com a mínima perda de conteúdo e funcionalidade” (NISO, 2004, p. 02, tradução nossa).

Com a evolução no desenvolvimento dos metadados, outros requisitos passaram a ser necessários, proporcionando o desenvolvimento de um conjunto mais abrangente e específico de princípios e práticas. A seguir serão explicadas as principais características dos princípios e práticas atuais.

Conforme salientam Duval et.al (2002), os princípios podem ser considerados “conceitos julgados comuns para todos os domínios de metadados e que possam subsidiar a formulação de qualquer esquema de metadados ou aplicação”. Eles orientam o desenvolvimento de soluções práticas em qualquer domínio em relação aos aspectos semânticos e de interoperabilidade dos padrões de metadados. Duval et.al (2002), destacam como princípios:

- *Modularidade*: permite que os desenvolvedores de padrões de metadados reutilizem os atributos em outras estruturas semânticas e sintáticas de padrões de metadados, ao invés de reinventar elementos. Refere-se à construção de metadados em blocos, categorias, grupos ou módulos de elementos descritivos, de modo que esses metadados possam ser estruturados em categorias de outros padrões de metadados com estruturas sintática e semanticamente diferentes, mas que possam ser interoperáveis. Em outras palavras, diferentes categorias semânticas de metadados são combinadas em outros padrões de metadados com estrutura sintática diferente, porém expressas em uma linguagem comum, incorporando a funcionalidade e o significado de cada atributo. Dessa forma, conjuntos de módulos podem ser criados para atender a requisitos específicos de uma aplicação, combinando exigências específicas de um domínio com exigências gerais, sem danos à interoperabilidade (DUVAL et.al, 2002; ZENG; QIN, 2008).
- *Extensibilidade*: um dos princípios indicados pelo DCMI, conforme já explicado, é a capacidade de extensão, ou seja, de permitir a inclusão de novos metadados em um padrão de acordo com as necessidades específicas de aplicação (DUVAL et.al, 2002);

- *Refinamento*: as aplicações de metadados em um domínio irão variar de acordo com a necessidade do grau de detalhe ou especificidade na descrição; dessa forma, os padrões de metadados devem permitir a escolha do nível de detalhes na descrição dos recursos. Duas noções de refinamento devem ser consideradas: a adição de qualificadores, que são elementos de metadados que qualificam e tornam o significado de um atributo mais específico (semelhantes a subcampos de descrição); e esquemas de codificação externos ao padrão de metadados utilizados para padronizar a representação determinando os valores do elemento, como, por exemplo, vocabulários controlados, esquemas de codificação do conteúdo, notações de sistemas de classificação etc. (HILLMANN, 2005; DUVAL et.al, 2002; ZENG; QIN, 2008).
- *Multilinguismo*: incide sobre os aspectos da diversidade cultural e linguística que devem ser considerados no momento de estabelecer os metadados de um padrão. Por exemplo, a maneira como as datas são representadas em diferentes calendários; a direção na qual o texto é apresentado e lido; conotações culturais de certos ícones e pictogramas, entre outros. É importante que os metadados descrevam características relevantes de um recurso respeitando as diferenças culturais e linguísticas (DUVAL et.al, 2002; ZENG; QIN, 2008).

Já as práticas, de acordo com Duval et.al (2002), podem ser consideradas “regras de manuseio, restrições, e problemas de infraestrutura que emergem por trazer a teoria em prática sob a forma de sistemas úteis e sustentáveis”. As práticas referem-se aos aspectos emergentes dos metadados em relação a sua criação e gerenciamento. Duval et.al (2002) destacam como práticas:

- *Perfil de aplicação*: é um conjunto de elementos de metadados selecionados de um ou mais esquema de metadados, combinados em um esquema composto. Os perfis de aplicação oferecem meios para se expressar os princípios de modularidade e extensibilidade. Sua finalidade é combinar esquemas de metadados adaptando-os às exigências funcionais de aplicações específicas, e mantendo a interoperabilidade com esquemas originais que serviram como base. Seu maior objetivo é facilitar a interoperabilidade semântica (DUVAL et.al, 2002).
- *Sintaxe e semântica*: é necessário um acordo comum entre dois domínios, tanto a sintaxe quanto a semântica dos metadados. Em outras palavras, é preciso que domínios diferentes compartilhem o mesmo significado de um atributo ou metadado e que estabeleçam uma

convenção comum para a identificação e a codificação dos valores em uma sintaxe, caso contrário não haverá intercâmbio de dados (DUVAL et.al, 2002).

- *Associações entre modelos*: referem-se à associação entre diferentes registros de metadados de um mesmo recurso informacional. Múltiplos registros de metadados refletem os vários propósitos e perspectivas de diferentes organizações ao criarem e gerenciarem metadados. Espera-se que esses registros sejam gerenciados de forma coordenada (DUVAL et.al, 2002).
- *Identificando e nomeando elementos de metadados: símbolos X etiquetas*: os metadados são expressos por um nome que pode ser codificado por meio de uma etiqueta em linguagem natural ou codificado em símbolos. Ferramentas específicas podem utilizar rótulos de apresentação diferentes da codificação original dos metadados para melhor visualização e entendimento (DUVAL et.al, 2002).
- *Registro de metadados*: um sistema formal para documentação de conjuntos de metadados, perfil de aplicação, esquemas de codificação, informações de uso do elemento e *crosswalks*. A função primária dos registros de metadados é registrar, publicar e gerenciar essas especificações, como também facilitar a busca dentro do registro (ZENG, QIN, 2008, p. 323).
- *Integralidade na descrição*: nem todos os atributos disponíveis em um esquema de metadados serão utilizados ou são adequados para representar um recurso; a representação deve conter os atributos essenciais para caracterizar o recurso. Além disso, a seleção apropriada dos valores também se constitui em um item importante para compor a totalidade da descrição. Assim, a riqueza na descrição será determinada por políticas de melhores práticas estabelecidas pela agência criadora dos metadados e construídas a partir de requisitos funcionais ou serviços de aplicação em diferentes domínios (DUVAL et.al, 2002). Políticas de melhores práticas são orientações para a realização de uma tarefa de modo mais eficiente (com menos esforço) e mais efetivo (com melhores resultados). Não se constituem nem como um padrão nem como um regulamento, mas sim como diretrizes a serem seguidas para enriquecer a representação (ZENG, QIN, 2008).
- *Elementos obrigatórios X opcionais*: padrões de metadados para propósitos gerais requerem um alto grau de flexibilidade. Dessa forma, um elemento de metadado pode ser essencial em um domínio e em outro não, por não refletir as características de um determinado recurso. Por esse motivo, não há com pensar em obrigatoriedade em algum elemento de metadados. Contudo, é importante se pensar em políticas de boas práticas

para especificar normas de conduta para o uso de um determinado padrão em um domínio, favorecendo, assim, a padronização na estrutura dos elementos em um domínio e consequentemente a interoperabilidade (DUVAL et.al, 2002).

- *Metadados subjetivos X objetivos*: o processo de criação de metadados pode envolver metadados objetivos, aqueles que são características de fato, tais como título, autor, data; ou metadados subjetivos, aqueles que são atribuídos, tais como assunto, palavras-chave, resumo etc. Outros metadados são subjetivos por estarem sujeitos a interpretações sob diferentes pontos de vista ou domínios. Dessa forma, um dos requisitos no desenvolvimento de metadados é explicitar o contexto da informação, para que os aplicativos possam identificar mais facilmente o contexto onde está inserido o recurso (DUVAL et.al, 2002).
- *Geração automática de metadados*: entre a criação manual de representações em padrões de metadados e entre as ferramentas de *Harvesting* e indexação automática de metadados, existe uma variedade de outras aplicações de padrões as quais podem ser utilizadas para a geração automática de metadados. É possível reduzir custos com o processamento automático das informações na medida em que as descrições e o estabelecimento de metadados sejam mais consistentes (DUVAL et.al, 2002).

Conforme se percebe, esses princípios e práticas se configuram em um panorama mais geral. Existem ainda outros princípios, específicos de cada domínio, e ainda os princípios que norteiam a criação de cada padrão de metadados.

Tanto os princípios, quanto os padrões, como também os outros instrumentos utilizados para a representação informacional devem estar em consonância para proporcionar a efetividade dos sistemas e a interoperabilidade.

Após uma revisão dos fundamentos, conceitos, tipologia, funções, características, aplicações e princípios dos metadados, serão abordados no próximo capítulo os padrões de metadados para propósito geral (padrões para a Web), os padrões de metadados de um domínio específico (padrões para o domínio bibliográfico), os princípios que nortearam sua criação e demais características relevantes para essa pesquisa.

4 PADRÕES DE METADADOS

Os metadados estão presentes nos sistemas sob diferentes tipos, níveis de especificidade, estrutura e aprimoramento, conforme pode ser observado pelos diferentes padrões de metadados existentes (MÉNDEZ RODRÍGUEZ, 2002).

A necessidade de representar as informações existe em diversas áreas do conhecimento, e para atender a essa necessidade foram sendo criados vários padrões de metadados que variam desde estruturas simples, passando por um tipo de padrão intermediário, até padrões de estruturas mais complexas de descrição. O uso do padrão de metadados correspondente ao tipo específico de ambiente informacional, bem como a construção adequada e padronizada de representações se constitui na chave para a efetividade dos sistemas de informação nesses ambientes.

A aplicação apropriada do padrão de metadados garante uma descrição normalizada e como consequência uma representação de qualidade, facilitando o intercâmbio de informações, a interoperabilidade entre sistemas e a recuperação da informação.

Entretanto, o que determina a aplicação de um padrão de metadados é o princípio pelo qual ele foi criado. Isso significa que o princípio norteará toda a característica, a finalidade e o uso do padrão e conseqüentemente a representação gerada a partir de sua estrutura descritiva (conjunto de metadados). Assim, quanto mais específico for o ambiente informacional, maior será a exigência de especificidade na descrição e, portanto, deverá ser utilizado um padrão de metadados correspondente a essa necessidade. Caso isso não seja feito, a recuperação da informação será insuficiente.

Deste modo, para o uso e a aplicação adequada dos padrões de metadados é necessário conhecer as particularidades e diferenças entre suas estruturas e níveis de especificidade, e principalmente conhecer o princípio que norteou sua criação, ou seja, a finalidade para a qual foi desenvolvido. Essas características determinam não só a aplicação dos padrões de metadados como também direcionam o seu uso, de tal forma que a recuperação baseada em metadados possa ser realizada com maior qualidade.

Segundo os autores Dempsey e Heery (1997), os formatos ou padrões de metadados podem ser entendidos sob três tipologias ou níveis, conforme aponta o quadro a seguir:

	BANDA UM	BANDA DOIS	BANDA TRÊS
Característica do registro	Formatos simples	Formatos estruturados	Formatos ricos
	Padrão proprietário	Padrões emergentes	Padrões internacionais
	Indexação do texto completo	Estrutura em campos	Identificação detalhada (descrição em campos ou tags codificadas)
Formatos dos registros	Lycos, Altavista, Yahoo etc.	Dublin Core, IAFA Templates, RCF 1807, SOIF, LDIF	MARC, TEI, CIMI, EAD, ICPSR

QUADRO 3: Tipologia de formatos de metadados.
 FONTE: Dempsey; Heery (1997, p. 08, tradução nossa).

De acordo com as tipologias ou níveis, denominados no quadro anterior como Banda um, Banda dois e Banda três, é possível encontrar tipos diferentes de padrões de metadados que variam desde estruturas mais simples de descrição, passando por uma estrutura intermediária até chegarem em uma estrutura mais complexa de descrição. Esses tipos de padrões, também chamados de formatos de metadados, apresentam as seguintes características, conforme apontado nos estudos de Alves (2005):

- a) Formatos simples: constituídos por metadados não estruturados, extraídos de forma automática por robôs, apresentam na maioria das vezes uma semântica reduzida. Ex.: MetaTag(s) e metadados utilizados na transferência de dados por meio do protocolo HTTP - hipertext transfer protocol.
- b) Formatos estruturados: constituídos por metadados mais estruturados, baseados em normas emergentes e que proporcionam uma descrição mínima do recurso para sua identificação, localização e recuperação. A descrição geralmente é feita em campos e nessa categoria começa a ser inserida a ajuda de especialistas em informação. Ex.: padrão de metadados Dublin Core – DC.
- c) Formatos ricos: também considerados padrões de metadados altamente estruturados, são constituídos por metadados complexos, apresentam uma estrutura de descrição mais formal e detalhada. São baseados em normas e códigos especializados de um domínio particular, possibilitam a descrição de um recurso informacional individual ou pertencente a uma coleção e facilitam a localização, a recuperação, o intercâmbio dos recursos informacionais. Ex.: padrão de metadados ou formato MARC 21, da área de Biblioteconomia.

Além das características básicas apontadas anteriormente, é necessário saber que cada tipo de padrão apresenta uma funcionalidade e valor em ambientes informacionais digitais. No entanto, é importante que algumas considerações sejam traçadas para entender os motivos pelos quais o uso de padrões de metadados passou a ser considerado cada vez mais necessário.

Os padrões de metadados simples, como as MetaTag(s), URI (Uniform Resource Identifier) e os dados trocados na transferência do protocolo HTTP, são iniciativas importantes na tentativa de estabelecer a localização e a recuperação dos recursos informacionais na Web, porém insuficientes. As MetaTag(s) são os únicos elementos que comportam algum tipo de descrição dos recursos e se constituem na forma mais comum de uso de algum elemento descritivo na Web. O uso de MetaTag(s) descritivas auxilia a indexação feita pelas ferramentas de busca e conseqüentemente a recuperação, pois evita que o texto seja indexado na íntegra. Contudo, apresenta como desvantagem a falta de um controle formal na descrição, pois muitas vezes não são informadas nas Tag(s) as atualizações do recurso informacional. Por esse motivo, os padrões de metadados simples são considerados insuficientes quando se trata de propor uma descrição mais formal do recurso que possa auxiliar efetivamente na recuperação no mesmo.

Os padrões de metadados estruturados constituem-se como uma tendência atual a ser utilizada na Web, no intuito de oferecer a descrição estruturada dos recursos informacionais. O padrão de metadados Dublin Core – DC – é um exemplo; foi originalmente criado para promover a descoberta de recursos informacionais na Web por meio de uma descrição e identificação mínima (WOODLEY, CLEMENT, WINN, 2005). Além disso, vem ganhando destaque em diversos ambientes informacionais devido a sua simplicidade, flexibilidade e abrangência. Sendo assim, o princípio de criação ou a razão da origem do padrão é promover a descoberta de recursos informacionais na Web. Esse princípio norteará as características da estrutura descritiva e conseqüentemente a representação dela originada.

Os padrões de metadados ricos ou altamente estruturados caracterizam-se por serem específicos de um domínio. Por mais que se busque uma maior funcionalidade e flexibilidade dos padrões de metadados por meio de estruturas simples de descrição, não é possível evitar a complexidade de descrição, criada pela necessidade de especificidade em um domínio. Por esse motivo, os padrões de metadados ricos apresentam uma estrutura de descrição formal, baseada em normas e códigos especializados, e um nível de descrição detalhado, ou seja, descrevem o recurso informacional com especificidade. O formato MARC 21 se constitui como um exemplo de formato ou padrão de metadados desta tipologia no domínio

bibliográfico. O formato MARC reflete o grau de desenvolvimento e especificidade da área de Biblioteconomia em relação à descrição de diversos tipos de recursos informacionais em uma unidade de informação.

É importante considerar que os três níveis abordados por Dempsey e Heery (1997) não são categorias estanques. Especialistas em informação, como indexadores e catalogadores, tipicamente presentes na Banda três, podem assumir algumas atividades específicas da Banda dois, tanto para divulgar seus recursos de forma mais ampla com formatos de metadados de outras categorias e ambientes informacionais, como também para incrementar a descrição de formatos característicos da Banda dois (ZENG, QIN, 2008). Isso demonstra que independente do nível de padrão de metadado e de seu objetivo em um determinado ambiente informacional, o profissional da informação vem ganhando cada vez mais espaço e importância em sua atuação.

Para esta pesquisa, os padrões a serem estudados pertencem à Banda dois – formatos estruturados e Banda três – formatos ricos, considerados também padrões de metadados para propósitos gerais e padrões de metadados para domínio específico, respectivamente (BARRETO, 1999; ALVES, 2005; ZENG, QIN, 2008).

Dessa forma, os formatos de metadados analisados serão o padrão de metadados Dublin Core, responsável pela descrição de recursos informacionais na Web (formato estruturado) e o padrão de metadados MARC 21, responsável pela descrição de recursos informacionais no domínio bibliográfico (formato rico).

4.1 PADRÕES DE METADADOS PARA PROPÓSITOS GERAIS: formatos estruturados para o domínio da Web

Os padrões de metadados para propósitos gerais são compostos pelos padrões de metadados simples e estruturados; contudo, serão abordados aqui somente os padrões de metadados estruturados criados especificamente para promover uma melhor recuperação no domínio da Web.

O padrão de metadados que mais se destaca nesta categoria é o Dublin Core, criado especificamente para o domínio da Web. Embora seja muito estudado na literatura, serão abordados nesta pesquisa seu histórico, de forma breve, e algumas de suas principais características, principalmente em relação ao seu escopo, seu princípio de criação e sua

estrutura de descrição. Antes de serem abordadas as características desse padrão, é preciso entender algumas denominações sobre os padrões de metadados, conforme explicam Zeng e Qin (2008).

Os formatos de metadados estão intrinsecamente ligados com a determinação dos metadados e vice-versa. A construção de metadados padronizados se efetivará a partir do estabelecimento de um formato de metadados; igualmente, o estabelecimento de um formato de metadados depende da criação de uma estrutura descritiva de metadados padronizados (conjunto de elementos de metadados - atributos). Essa relação se constitui em uma das chaves da efetivação dos sistemas. Sendo assim, pode-se dizer que os formatos de metadados foram criados para atender a um desígnio mais geral, que é guiar a construção padronizada de metadados por meio da implementação de uma estrutura de dados, valores, conteúdo e intercâmbio de dados de modo eficiente e consistente (ZENG, QIN, 2008).

Desta forma, pode-se dizer que um formato ou padrão de metadados é constituído basicamente dos seguintes componentes:

- Padrões para estrutura de dados: também denominados Esquema de metadados (*Metadata Scheme*), são formados pelo conjunto de elementos de metadados (*element sets*) metodologicamente construídos e uniformizados segundo padrões de estrutura de dados e padrões semânticos. “[...] Este tipo de padrão especifica o que constitui um registro e define as categorias de informação ou campos de dados, assim como as relações entre eles. [...]” (BACA, 1998, p. 37, tradução nossa). Os padrões de estrutura estabelecem regras para a padronização dos elementos descritivos ou metadados em um formato, definem as características dos campos descritivos e as relações entre os elementos (ZENG, QIN, 2008; MÉNDEZ RODRÍGUEZ, 2002).
- Padrões para conteúdo de dados: referem-se aos padrões externos que guiam a ordem, a sintaxe e a forma dos valores dos metadados nos padrões de metadados. Esses padrões são indicados nos esquemas de codificação utilizados nos padrões de metadados (ZENG, QIN, 2008; MÉNDEZ RODRÍGUEZ, 2002). Serão abordados mais especificamente no item 5.2.
- Padrões para valores dos dados: também se constituem como padrões indicados nos esquemas de codificação dos padrões e determinam os valores dos dados em um padrão de metadados (ZENG, QIN, 2008; MÉNDEZ RODRÍGUEZ, 2002). Esse tipo de padrão também será abordado de modo específico no item 5.2.

- Padrões para intercâmbio de dados: referem-se aos diferentes formatos, linguagem de estruturação dos dados ou protocolos utilizados para a comunicação e intercâmbio dos dados em um sistema. Podem ser projetados juntamente com a estrutura de dados do padrão de metadados ou separadamente (ZENG, QIN, 2008). Também será abordado de modo específico no item 5.2.

Esses componentes estarão presentes nos formatos de metadados estruturados e nos formatos de metadados ricos, porém com um nível maior de especificidade. A seguir serão abordadas as características do padrão de metadados Dublin Core específico da categoria de formatos estruturados.

O padrão de metadados Dublin Core - DC - se constitui em um simples e eficaz esquema de metadados utilizado para descrever uma ampla variedade de recursos em diversos ambientes informacionais da Web. Foi criado em 1995 por profissionais de diversas áreas do conhecimento, com o intuito de promover a descoberta de recursos na Web; atualmente vem ganhando destaque devido a sua simplicidade e flexibilidade para descrição de recursos informacionais (HILLMANN, 2005; GRÁCIO, 2002).

Estruturalmente, o padrão de metadados DC se constitui em dois níveis: simples e qualificado.

- a) Nível simples: contempla quinze elementos de descrição;
- b) Nível qualificado: contempla os quinze elementos de descrição, acrescidos de sete elementos adicionais e duas classes de qualificadores:
 - Elementos de refinamento: são qualificadores utilizados com o intuito de refinar ou restringir, especificando e ampliando o significado dos elementos descritivos;
 - Esquemas de codificação: qualificadores que identificam esquemas para o valor do elemento, como vocabulário controlado, notações de sistemas de classificação, entre outros (HILLMANN, 2005).

O quadro a seguir mostra os elementos descritivos componentes do Esquema de Metadados do padrão DC. Os primeiros quinze elementos pertencem ao Nível simples, enquanto que todos os elementos apresentados no quadro abaixo pertencem ao Nível qualificado, bem como os elementos de refinamento e esquemas de codificação. Optou-se por não traduzir os elementos, já que originalmente os elementos descritivos se apresentam na língua inglesa.

ELEMENTOS DO ESQUEMA DE METADADOS DC	ELEMENTOS DE REFINAMENTO		ESQUEMAS DE CODIFICAÇÃO	
Title	Alternative		-	
Creator	-		-	
Subject	-		LCSH MeSH DDC	LCC UDC
Description	Table Of Contents Abstract		-	
Publisher	-		-	
Contributor	-		-	
Date	Created Valid Available Issued	Modified Date Accepted Date Copyrighted Date Submitted	DCMI Period / W3C-DTF	
Type	-		DCMI Type Vocabulary	
Format	-		IMT	
	Extent		-	
	Medium		-	
Identifier	-		URI	
	Bibliographic Citation		-	
Source	-		URI	
Language	-		ISO 639-2RFC 3066	
Relation	Is Version Of Has Version Is Replaced By Replaces Is Required By Requires Is Part Of	Has Part Is Referenced By References Is Format Of Has Format Conforms To	URI	
Coverage	Spatial		DCMI Point ISO 3166	DCMI Box TGN
	Temporal		DCMI Period W3C-DTF	
Rights	Access Rights		-	
	License		URI	
ELEMENTOS ADICIONAIS	ELEMENTOS DE REFINAMENTO		ESQUEMAS DE CODIFICAÇÃO	
Audience	Mediator Education Level		-	
Provenance	-		-	
Rights Holder	-		-	
Instructional Method	-		-	
Accrual Method	-		-	
Accrual Periodicity	-		-	
Accrual Policy	-		-	

QUADRO4: Elementos do Esquema de Metadados DC.
FONTE: HILLMANN (2005).

Tanto os quinze elementos de descrição como os elementos adicionais, elementos de refinamento e esquemas são descritos em uma estrutura de codificação simples em forma de pares como as Tags: <nome=atributo propriedade=valor>. Ex.: <meta name="DC.Title" content="National Library of Australia">.

Todos os elementos são opcionais e podem ser repetidos, demonstrando a característica de flexibilidade do padrão, além disso, o nível de especificidade proporcionado pelos elementos de refinamento e pelos esquemas de codificação também é opcional. Desta forma, existirá uma diversidade de estruturas descritivas com diferentes conjuntos de metadados de acordo com a necessidade de cada ambiente informacional. Cabe ao indivíduo responsável pela representação do recurso decidir qual estrutura descritiva adotará.

Os autores Woodley, Clement e Winn (2005) apontam que o padrão DC foi originalmente criado para promover a descoberta de recursos informacionais na Web por meio de uma descrição e identificação mínima. Dessa forma, o princípio de criação ou a razão da origem do padrão é promover a descoberta de recursos informacionais na Web. Esse princípio norteará as características da estrutura descritiva e consequentemente a representação dela originada.

Duas categorias de princípios estão relacionadas a este padrão: princípios relacionados aos metadados e aos recursos que eles descrevem, e os princípios gerais para o estabelecimento da estrutura do esquema de metadados.

Em relação à primeira categoria, princípios relacionados aos metadados e aos recursos que eles descrevem, existem três princípios básicos que relacionam os metadados com os recursos:

- Princípio um para um: em geral os metadados do padrão DC descrevem uma manifestação ou uma versão de um recurso (um recurso corresponde a uma representação), ao invés de relacionar manifestações. Por exemplo, uma reprodução de uma imagem não terá em sua descrição pelo padrão DC uma relação com a imagem original. Cada uma apresentará uma descrição diferente e não relacionada (HILLMANN, 2005).
- Princípio “Dumb-down”: esse princípio determina que o implementador do padrão seja capaz de ignorar qualquer qualificador e de utilizar o valor como se fosse absoluto, pois embora isso signifique perda de especificidade, o valor do elemento sem o qualificador deve ser válido e útil para descoberta. O qualificador só detalha o significado semântico do elemento de metadado (HILLMANN, 2005).
- Valor apropriado: refere-se às boas práticas para um determinado elemento ou qualificador em contextos variados. De modo geral, um elemento deve manter seu valor correspondente como condição para a descoberta do recurso (HILLMANN, 2005).

Quanto aos princípios gerais para o estabelecimento da estrutura do esquema de metadados, Hillmann (2005) aponta que o padrão DC apresenta as seguintes metas:

- Simplicidade de descrição e manutenção: o esquema de metadados DC foi criado com o intuito de apresentar um conjunto mínimo de elementos descritivos e simples, para que pudesse ser facilmente descrito e gerenciado por não especialistas em informação, ao mesmo tempo em que garante a recuperação dos recursos por meio da descrição mínima (HILLMANN, 2005).
- Semântica comumente compreendida: a descoberta de informação é dificultada pelas diferenças terminológicas e práticas descritivas entre áreas ou domínios do conhecimento. O padrão DC fornece um esquema de metadados genérico, mas com semântica universalmente compreendida, esse princípio garante que a interoperabilidade semântica seja realizada em diferentes domínios do conhecimento (HILLMANN, 2005).
- Alcance Internacional: refere-se à participação de vários países no DCMI (*Dublin Core Metadata Initiative*) na busca de um escopo internacional para o estabelecimento de uma infraestrutura para o padrão DC em relação a um consenso sobre aspectos de natureza multilíngue e multicultural da informação em meio digital (HILLMANN, 2005).
- Extensibilidade: embora seja um padrão simples, o DC segue o princípio de extensibilidade, ou seja, a capacidade de estender e acrescentar outros elementos descritivos ao esquema de metadados do padrão. Dessa forma, é possível estender o conjunto de elementos com metadados adicionais específicos de um domínio (perfis de aplicação) (HILLMANN, 2005).

É importante destacar que a semântica estabelecida no padrão (o significado dos metadados e seus elementos de refinamento) foi projetada para ser independente de sintaxe, ou seja, o padrão estabelece uma estrutura e uma semântica para os dados, mas a sintaxe (linguagem) de codificação é estabelecida de modo diferente em cada domínio. Isso facilita o desenvolvimento de mapeamentos e correspondências entre diferentes sintaxes de domínios de aplicações distintos. Como exemplo de sintaxes utilizadas citam-se: linguagem HTML, linguagem XHTML, linguagem XML, RDF / XML entre outras. O DCMI vem desenvolvendo e disponibilizando diretrizes para a codificação do padrão DC em diferentes sintaxes (HILLMANN, 2005).

Além disso, o uso de esquemas de codificação vem sendo recomendado para proporcionar uma melhor representação dos recursos na Web. Conforme mencionado no capítulo anterior, o uso de esquemas de codificação nos elementos de refinamento do padrão se constitui em um dos princípios gerais recomendados.

Apesar de o padrão de metadados DC estar em consonância com princípios internacionais, é preciso destacar que ele apresenta certa limitação quando se refere à

descrição de acervos específicos disponíveis em ambiente Web, pois apresenta uma estrutura descritiva simples e genérica para abranger a maior variedade de tipos de recursos informacionais e, portanto, não possui a especificidade necessária de um domínio.

Muito embora os desenvolvedores do padrão DC estejam trabalhando para desenvolver meios de equilibrar a questão da simplicidade na descrição, com a necessidade de precisão na recuperação por meio de extensões em um domínio específico por meio de perfil de aplicação (HILLMANN, 2005), é preciso considerar que em muitas áreas ou domínios a construção de metadados já vem sendo realizada há muito tempo. É o que ocorre na área de Biblioteconomia, por exemplo, na qual encontramos formatos de metadados ricos, abordados a seguir.

4.2 PADRÕES DE METADADOS PARA PROPÓSITOS ESPECÍFICOS: formatos ricos para o domínio bibliográfico

Os padrões de metadados para propósitos específicos são compostos pelos padrões de metadados ricos ou altamente estruturados, apresentam um esquema de metadados complexo que reflete o grau de especificidade descritiva de um domínio específico. O padrão de metadados abordado aqui será o formato MARC 21 (*MAchine-Readable Cataloguing*) específico da área de Biblioteconomia e o mais utilizado no domínio bibliográfico. Assim, será abordado aqui um breve panorama do formato, destacando suas principais características no que se refere ao seu escopo, princípio de criação, estrutura de descrição e sua relação com os princípios de metadados.

De acordo com Chan (2007, p. 447, tradução nossa), pode-se afirmar que

O formato MARC 21 é um conjunto de padrões desenvolvidos com a finalidade de representar e comunicar de forma legível por máquina, metadados descritivos sobre itens de informação – particularmente, mas não somente, itens bibliográficos.

Foi desenvolvido nos anos de 1960 pela Library of Congress (LC) com o intuito de ser um formato padronizado para a descrição, armazenamento e intercâmbio automatizado de registros bibliográficos e catalográficos de diferentes tipos de recursos informacionais em uma unidade de informação (BRITISH LIBRARY, 2003).

Não é objetivo deste capítulo abordar em detalhes o histórico de desenvolvimento e evolução do formato MARC, pois isso já foi feito por diversos autores, como Chan (2007),

que trata de forma breve e clara o histórico do formato. Porém, é importante ressaltar que o histórico de desenvolvimento do formato MARC 21 reflete o amadurecimento da área em relação ao aprimoramento das regras de catalogação em uma estrutura automatizada. Sendo assim, não só a estrutura descritiva do formato MARC passou por inúmeras atualizações e unificações, como também as regras de representação correspondentes.

Com base nas afirmações acima, pode-se considerar que o princípio que norteou o desenvolvimento do formato MARC 21 foi a criação de uma estrutura descritiva padronizada e internacionalmente aceita pela comunidade biblioteconômica, para que pudesse representar e intercambiar dados bibliográficos de modo automatizado.

Conforme apontado por Zeng e Qin (2008) existe uma intrínseca ligação no estabelecimento de metadados e formatos de metadados. O estabelecimento de um formato de metadados exige a construção padronizada de metadados (atributos que representarão uma entidade), assim como os metadados precisam estar em uma estrutura de descrição padronizada (esquema de metadados) que irá compor o formato para serem úteis e válidos em um sistema de informação.

Uma grande diferença entre os padrões de metadados da categoria anterior e desta refere-se à criação do esquema de metadados. Na categoria anterior, o esquema de metadados foi sendo desenvolvido juntamente com o estabelecimento dos metadados para possibilitar a construção padronizada de representações para o domínio da Web. Com base nos princípios gerais já abordados, foram sendo estabelecidos: padrões para a estrutura de dados (ou esquema de metadados), padrões para dados de conteúdo, padrões para os valores dos dados e padrões para intercâmbio dos dados.

Já no domínio bibliográfico, os metadados ou os atributos que representam uma entidade já haviam sido criados, aperfeiçoados e padronizados ao longo do tempo, de acordo com princípios, códigos e regras de catalogação. Dessa forma, pode-se considerar que o esquema de metadados no domínio bibliográfico estava parcialmente determinado; com o avanço tecnológico era preciso adaptar as regras e metadados existentes em uma estrutura de dados ou esquema de metadados de modo automatizado e, assim, garantir o intercâmbio de dados por meio do estabelecimento de padrões.

Assim, os padrões de conteúdo de dados, padrões externos, códigos e regras que auxiliam a descrição e, conseqüentemente, o estabelecimento dos metadados, determinam não só a sintaxe e a padronização dos metadados, mas também os valores e padronizam sua representação. Os padrões de valores de dados, que determinam os valores semânticos dos dados em um padrão de metadados, são instrumentos criados e utilizados há muito tempo no

domínio bibliográfico e se constituem como instrumentos essenciais para o estabelecimento de metadados e formatos de metadados neste domínio (ZENG, QIN, 2008).

Pode-se considerar que o esquema de metadados do domínio bibliográfico, representado pelo formato de metadados MARC 21, foi estabelecido com base na codificação de metadados existentes no domínio e que foram metodologicamente construídos e uniformizados por princípios, códigos e regras de catalogação. O estabelecimento do formato MARC 21 possibilitou a existência de uma estrutura de dados (esquema de metadados) que refletisse os metadados específicos do domínio bibliográfico, que garantisse uma representação formal, padronizada e de alto nível de descrição, bem como facilitasse o intercâmbio dos registros bibliográficos criados por meio da estrutura de dados.

Embora na época não fossem utilizados os termos metadados e padrões de metadados, pode-se considerar que o formato MARC 21 se constitui, sem dúvida, no padrão de metadados do domínio bibliográfico. O formato MARC 21 apresenta como escopo criar representações padronizadas dos recursos informacionais de uma unidade de informação, por meio de seu esquema de metadados e esquemas associados (princípios e códigos de catalogação), com o intuito de intercambiar metadados descritivos ou mais especificamente metadados bibliográficos.

Assim, os metadados que compõem a estrutura do formato MARC 21 representam bibliograficamente um recurso informacional de modo detalhado e completo, refletindo no meio tecnológico as características específicas de descrição advindas do processo de catalogação. Dessa forma, pode-se considerar também que o formato MARC 21 reflete o grau de desenvolvimento e especificidade da área em relação à descrição de diversos tipos de recursos informacionais em uma unidade de informação, e principalmente, demonstra a integração estratégica dos tradicionais princípios de tratamento descritivo da informação (TDI) com as vantagens trazidas pelo uso de tecnologias de informática.

Quanto à estrutura, o formato MARC 21 é composto por cinco tipos de padrões, denominados também formatos (CHAN, 2007). Esse conjunto de padrões compõe o que se denomina família MARC, apresentada a seguir:

- MARC21 para dados bibliográficos (bibliographic data): 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);
- MARC21 para dados de autoridade (authority data): 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 utilizadas como pontos de acesso no registro bibliográfico;

- MARC21 Holdings (holdings data – controle de dados): contém especificações para a codificação de elementos identificados por um código no Líder (posição 06), informação utilizada para o gerenciamento de recursos;
- MARC21 para dados de classificação (classification data): contém especificações para a codificação de elementos relacionados a esquemas de classificação;
- MARC21 para informação de comunidade (community information): contém especificações para a codificação de informações sobre recursos não bibliográficos (indivíduos, organizações, serviços etc.) destinados a uma comunidade específica (LIBRARY OF CONGRESS, 2006).

De acordo com Chan (2007), os quatro primeiros tipos são destinados à descrição de informação bibliográfica, o último destina-se a estruturar informações sobre indivíduos, organizações, programas, serviços, eventos, entre outras informações.

A característica sintática da estrutura do formato MARC 21 é comum para os cinco tipos de padrões citados anteriormente: a estrutura descritiva é complexa, apresenta metadados divididos em diretórios de dados, sendo cada elemento de metadados codificado em campos e subcampos representados simbolicamente por números. Relacionando com os estudos de Chen (1990), os campos e subcampos do MARC 21 são atributos artificiais codificados para representar uma característica da entidade ou recurso, ao invés de ser utilizado o nome do atributo como identificador da entidade na descrição.

Por ser um formato rico, está em consonância com princípios e regras específicas do domínio ao qual pertence. Assim, para padronizar a representação produzida pelo formato utiliza-se de princípios, códigos e regras de catalogação, tabelas de classificação, vocabulários controlados e outros instrumentos externos ao padrão para a construção de registros bibliográficos.

Quanto à especificidade, estrutura e aprimoramento, o formato MARC apresenta em sua estrutura de descrição todos os atributos necessários para representar diferentes tipos de recursos informacionais de modo detalhado, exaustivo para que possa identificá-lo de forma absoluta e única. Além disso, contém metadados ricos semanticamente, ou seja, o significado dos atributos que representam uma entidade ou recurso corresponde semanticamente às características dessa entidade (ZENG, QIN, 2008).

Embora os cinco formatos citados apresentem a mesma característica sintática, cada formato corresponde a um propósito de descrição; assim, cada um contemplará uma estrutura

de dados ou um esquema de metadados diferente. Em outras palavras, cada formato apresentará um conjunto diferente de metadados utilizados para representação.

Para esta pesquisa serão destacadas as características específicas do formato MARC 21 para Dados Bibliográficos, responsável por proporcionar a descrição bibliográfica de diferentes recursos informacionais, além de ser o mais conhecido e amplamente utilizado por bibliotecários.

Pode-se verificar a seguinte estrutura para o formato MARC 21:

- Estrutura do registro: “[...] é uma implementação da *American National Standard* para o Intercâmbio de Informação Bibliográfica (ANSI / NISO Z39.2) e a sua ISO equivalente ISO 2709.” (FERREIRA, 2002, p. iii);
- Indicação de conteúdo: “os códigos e convenções estabelecidos explicitamente para identificar e caracterizar os dados dentro do registro e permitir sua manipulação [...]” (FERREIRA, 2002, p.iii). A indicação de conteúdo é feita em campos ou etiquetas do tipo campos de dados variáveis dentro do diretório do formato. Os campos representados pelas etiquetas codificadas se constituem nos elementos de descrição do formato, conforme apontado resumidamente a seguir:
 - Campo: unidades lógicas que representam um tipo de informação bibliográfica;
 - Etiqueta: identifica o campo e a informação a ser representada; é composta por um número de três dígitos. Algumas podem ser duplicadas;
 - Indicador: são as duas primeiras posições de caracteres após a etiqueta; sua função é definir com maior detalhe os dados descritos nos campos; cada posição apresenta um valor a ser processado pela máquina;
 - Subcampo: alguns campos são divididos em subcampos ou elementos descritivos que especificam ainda mais a informação a ser representada;
 - Código de subcampo: os subcampos são representados por códigos 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: refere-se ao conjunto de etiquetas, indicadores e códigos de subcampo que, segundo Furrie (2001) 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, 2001). O designador de conteúdo pode ser considerado um elemento complexo de representação na estrutura do formato MARC 21.

- Conteúdo dos elementos que compõem o registro: o conteúdo dos elementos descritivos do formato MARC é composto pela representação de informações, geralmente definidas por padrões externos 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).

Para visualização da estrutura do formato MARC, apresenta-se a seguir o campo título e indicação de responsabilidade, codificado pela etiqueta 245 apresentada em vermelho. Em azul, as duas posições dos indicadores; em laranja, um separador de subcampo e a letra minúscula “a” correspondente ao subcampo título; em verde, um separador de subcampo e a letra “b” correspondente ao subcampo complemento do título; na cor rosa, o separador de subcampo e a letra “c” correspondente à indicação de responsabilidade. As informações em preto correspondem ao conteúdo (valor) do campo, normalizado por padrões externos, neste exemplo o código de catalogação AACR2.

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

QUADRO 5: Exemplo do campo Título e indicação de responsabilidade, do formato MARC 21.
FONTE: do autor.

A evolução nas versões do formato MARC possibilitou a inclusão do campo 856 para a indicação da URL (Uniform Resource Locator) de recursos digitais no formato MARC 21 para dados bibliográficos, e seu esquema de metadados passou a compreender a descrição da localização de recursos e meio digital.

A estrutura do formato MARC 21 possibilita a construção de registros bibliográficos de modo que o mesmo represente uma variedade de tipos de recursos informacionais no que diz respeito ao seu conteúdo e forma; facilita a recuperação dos recursos em sistemas informacionais específicos, como os catálogos de bibliotecas; e promove o intercâmbio e compartilhamento de registros bibliográficos entre bibliotecas. É importante destacar que

[...] este intercâmbio de informações ocorre com o uso de protocolos, e não abrange somente a descrição de dados bibliográficos, mas também a representação sobre controle (gerenciamento) dos dados, dados relacionados à autoridade ou responsabilidade das obras, a classificação de documentos descritos e dados locais que podem ser incluídos nos catálogos para a

comunidade informacional onde o material está localizado (ALVES, 2005, p. 143).

O formato MARC 21 apresenta uma estrutura rígida de descrição que acaba sendo uma vantagem para promover adequadamente o intercâmbio de registros bibliográficos em redes de bibliotecas.

Entretanto, para a aplicação do padrão de metadados específico do domínio bibliográfico em ambiente Web, é necessário que o formato apresente uma sintaxe mais flexível para codificação dos dados, garantindo o intercâmbio dos dados e interoperabilidade entre sistemas em ambientes generalizados de troca de dados (ZENG, QIN, 2008).

A versão do MARC 21 na linguagem XML (*eXtensible Markup Language*) traz benefícios não só para a comunidade biblioteconômica, mas também para outros domínios envolvidos com a criação e o estabelecimento de metadados e padrões de metadados (ZENG, QIN, 2008).

O MARC XML trouxe maior flexibilidade ao formato MARC 21, outro requisito importante nos princípios gerais para a construção de padrões de metadados atualmente. Dessa forma, bibliotecas digitais que compartilham metadados mesmo que em formatos distintos (padrão DC e padrão MARC 21, por exemplo), conseguem promover mais adequadamente a interoperabilidade dos dados, por compartilharem metadados em uma sintaxe de codificação compatível.

A proposta do MARC XML é solucionar a necessidade de flexibilidade, extensibilidade e modularidade sem, contudo, perder a especificidade de domínio, garantindo que a interoperabilidade entre sistemas não seja prejudicada.

Dentre as principais vantagens do MARC na versão XML estão: maior flexibilidade e extensibilidade do padrão; promoção do princípio de modularidade; garantia da interoperabilidade; expressão sintática do esquema de metadados do formato mantendo especificidade de domínio; os campos ou metadados codificados passam a ser representados em forma de tags; o XML *schema* mantém a semântica dos elementos de metadados do formato MARC; não há perda de informações no momento da conversão entre formato MARC 21 e formato MARC XML; a forma de apresentação da representação é feita por meio de uma folha de estilo associada à linguagem XML; pode ser utilizado para a representação de qualquer tipo de recurso informacional, entre outras vantagens (LIBRARY OF CONGRESS, 2002).

Com base nas principais características do formato MARC 21, seu escopo, princípio de criação, estrutura de descrição e sua relação com os princípios gerais dos metadados, é

preciso destacar aqui algumas considerações em relação aos formatos do domínio bibliográfico e aos formatos do domínio Web, respectivamente MARC 21 e DC.

Devido à variedade de padrões de metadados, é possível encontrar uma diversidade de tipos diferentes de representações de um mesmo recurso informacional. Quanto mais detalhada for a descrição, mais completa será a representação; quanto mais específica for a descrição, mais metadados ou atributos existirão para individualizar e caracterizar a entidade de forma inequívoca e única, garantindo, assim, maior quantidade de pontos de acesso ao recurso informacional e conseqüentemente, uma possibilidade de busca e recuperação mais eficiente.

Entretanto, é preciso levar em consideração o princípio de criação e o objetivo de cada padrão de metadados, pois isso definirá o esquema de metadados de cada padrão e a representação gerada por sua estrutura. O padrão de metadados DC não apresenta especificidade descritiva suficiente para representar os tipos de recursos de uma unidade de informação, enquanto que o formato MARC 21 apresenta uma estrutura descritiva complexa para ser utilizada por indivíduos que não tenham conhecimento especializado da área de Biblioteconomia.

Quando se trata de um domínio específico ou quanto mais especializado for o ambiente informacional, maior será a necessidade de descrição detalhada e maior será a necessidade de uso de um padrão de metadados altamente estruturado para garantir a especificidade de descrição necessária ao ambiente. Neste caso, a presença de especialistas no tratamento descritivo da informação referente a um determinado domínio é indispensável para garantir uma descrição concisa e padronizada do recurso.

Contrariamente, quanto mais abrangente for o ambiente informacional, maior a necessidade de um padrão de metadados menos estruturado, que contenha apenas descrições mais gerais e que possa ser utilizado por indivíduos não especialistas em tratamento informacional.

Portanto, ressalta-se a importância do conhecimento sobre o propósito de uso do padrão de metadados, ou seja, é importante saber se o princípio pelo qual o padrão de metadados foi criado corresponde ao tipo de ambiente informacional no qual será aplicado, para que possa fornecer uma representação adequada dos recursos e uma recuperação eficiente.

Atualmente presencia-se uma mudança no tratamento descritivo da informação, causada pelo uso de padrões de metadados. Essas mudanças estão ocorrendo não só na

concepção dos sistemas, mas se estendem principalmente no âmbito do tratamento descritivo da informação (MÉNDEZ RODRÍGUEZ , 2002).

Dessa forma, identificam-se atualmente duas tendências que se refletem nos sistemas de informação em diferentes domínios e ambientes digitais, relacionadas principalmente com os processos de recuperação da informação:

a) tendência tecnológica: protagonizada pelo desenvolvimento e aplicação de ferramentas tecnológicas que melhorem os processos de indexação; a aplicação da linguagem natural; o desenvolvimento de novos algoritmos de busca e recuperação; a visibilidade dos recursos informacionais por meio de linguagens de marcação melhor desenvolvidas e aplicadas; o desenvolvimento e aplicação de modelos mais consistentes para a construção de sistemas de bancos de dados etc (MÉNDEZ RODRÍGUEZ , 2002);

b) tendência representacional: protagonizada pelo desenvolvimento, aplicação e uso de padrões de metadados, e pelo uso de esquemas de codificação consistentes que garantam a representação única do recurso informacional. Essa tendência proporciona maior visibilidade semântica dos dados e representações, e também exige uma revisão dos processos existentes de tratamento descritivo da informação para que estejam em consonância com os aspectos tecnológicos em desenvolvimento (MÉNDEZ RODRÍGUEZ , 2002).

Destaca-se neste momento a importância de se conhecer os aspectos que levaram à determinação dos metadados e padrões de metadados para posteriormente estabelecer a categorização dos instrumentos necessários para a construção de representações padronizadas.



5 ASPECTOS TECNOLÓGICOS E REPRESENTACIONAIS PARA A DETERMINAÇÃO PADRONIZADA DOS METADADOS

A construção de representações por formatos ou padrões de metadados é algo complexo, de acordo com Méndez Rodríguez (2002, p. 192, tradução nossa): “Nos metadados convergem todos os aspectos relacionados com a normalização documental e todos aqueles relacionados com padrões técnicos de infraestrutura de comunicações na rede”.

Dessa forma, percebem-se dois aspectos em relação à determinação dos metadados:

- a) sua contextualização nos sistemas de informação, em especial nos sistemas de banco de dados que posteriormente se constituirão como repositórios de metadados, e
- b) sua determinação e padronização no que se refere ao uso de metodologias e instrumentos para representação informacional, ou seja, um aspecto diretamente relacionado com o tratamento descritivo da informação.

Esses dois aspectos se convergem no intuito de consolidar uma estrutura consistente de representação para difusão, intercâmbio e interoperabilidade dos dados entre sistemas informacionais em diferentes ambientes digitais (MÉNDEZ RODRÍGUEZ, 2002).

Na realidade, os aspectos tecnológicos e representacionais estão relacionados entre si: os aspectos tecnológicos representam por meio das ferramentas tecnológicas os aspectos representacionais, mais precisamente os aspectos de tratamento descritivo da informação, que por sua vez só existem no sistema de informação por meio das ferramentas tecnológicas que os consolidam.

Sem a união desses dois aspectos, não é possível estabelecer metadados consistentes⁴ em um sistema de informação: os aspectos representacionais determinam a forma coerente e lógica de representação de um domínio ao qual o padrão de metadados pertence, enquanto os aspectos tecnológicos consolidam e viabilizam os aspectos representacionais no sistema.

Nos capítulos anteriores foram abordadas revisões dos aspectos mais importantes sobre os fundamentos, os conceitos, as tipologias, as funções, as características, as aplicações dos metadados e os princípios gerais dos metadados, bem como as principais características dos padrões de metadados e seus níveis, destacando-se os padrões para propósitos gerais e específicos.

⁴ A consistência na determinação de metadados é sinônimo de coerência na representação, ou seja, ausência de contradição, fato que é proporcionado por uma estrutura descritiva lógica.

Os padrões de metadados dos níveis estruturado e rico destacados nesta pesquisa (respectivamente representados pelos padrões DC e MARC 21), pertencem à categoria de metadados que desempenham papéis tradicionais de identificação e descrição da informação, busca, recuperação e localização de recursos; ainda que em muitos momentos possam exercer alguns papéis emergentes como possibilitar o acesso aos conteúdos.

Apesar de serem padrões já consolidados, ressalta-se aqui a importância de conhecer como os metadados são determinados em um esquema de metadados de um padrão, pois isso irá nortear o aprimoramento dos esquemas desses padrões.

Embora os padrões DC e MARC tenham sido criados em contextos e domínios diferentes, ambos apresentam aspectos em comum: aspectos tecnológicos e representacionais.

A comunidade de metadados vem contribuindo para a consolidação desses aspectos com o desenvolvimento e a implantação de padrões de metadados e outras ferramentas tecnológicas associadas. Por isso, abordam-se nesta pesquisa os aspectos considerados mais relevantes para a determinação dos metadados desenvolvidos nessa comunidade.

Entretanto, conforme apontam Zeng e Qin (2008) várias decisões sobre os metadados e padrões de metadados devem ser adotadas em sua implementação em um sistema, e as características do domínio irão influenciar diretamente a escolha e adoção do tipo de padrão de metadados. Dessa forma, as peculiaridades do domínio serão preponderantes para determinar o padrão de metadados adequado a seu sistema.

Pode-se dizer que as características do domínio irão especificar a determinação dos metadados e padrões de metadados para um propósito particular, contudo a determinação estrutural dos metadados pode ocorrer de modo semelhante em alguns padrões, devido às características tecnológicas e representacionais que apresentam (ZENG, QIN, 2008).

Para uma implementação bem sucedida do sistema de informação, é importante conhecer o domínio, os aspectos tecnológicos e representacionais que norteiam a definição dos metadados e o estabelecimento deles (ZENG, QIN, 2008).

Dessa forma, o intuito deste capítulo é destacar os pontos mais importantes para a determinação padronizada dos metadados e padrões de metadados tendo como base os dois aspectos ressaltados: o aspecto tecnológico e o aspecto representacional.

5.1 ASPECTOS TECNOLÓGICOS PARA DETERMINAÇÃO DOS METADADOS: requisitos funcionais, modelagem conceitual e esquemas de metadados

A determinação dos metadados no que se refere ao aspecto tecnológico tem início na criação do sistema de informação, mais especificamente no projeto lógico da construção do sistema de banco de dados. Entretanto, em alguns casos a determinação dos metadados pode acontecer previamente no estabelecimento de um esquema descritivo teórico e depois passar a ser desenvolvido e aplicado em ambiente digital. Um exemplo a ser citado é o que ocorreu no domínio bibliográfico, no qual já apresentava um esquema descritivo por meio de seus códigos e regras, e posteriormente passou a existir em meio automatizado e digital com a estrutura de descrição do padrão MARC 21.

Independente do modo como isso ocorra, é preciso conhecer os aspectos envolvidos na criação dos metadados para desenvolver sistemas com maior eficiência.

Pode-se considerar que os aspectos tecnológicos para a determinação dos metadados estão diretamente relacionados com a modelagem conceitual, que vai determinar conceitualmente os metadados e os requisitos do sistema no banco de dados; e o estabelecimento de esquemas de metadados, que define a representação por meio de uma estrutura, sintaxe e semântica de codificação dos dados. A seguir, aborda-se cada deles.

5.1.1 REQUISITOS FUNCIONAIS E MODELAGEM CONCEITUAL

Os autores Zeng e Qin (2004, p. 87-88, tradução nossa) apontam que é necessário considerar que o projeto de uma coleção digital comece

[...] com a investigação sobre a disciplina específica, a comunidade de usuários potenciais e/ou usuários. Os fatores considerados incluem a natureza dos objetos da coleção; necessidades previstas dos usuários; e restrições sobre a criação de metadados, implementação, e controle de qualidade (recursos humanos, financeiro, equipamentos, interoperabilidade com serviços e sistemas institucionais e da comunidade de informação).

Com a análise desses aspectos, é possível definir: os princípios que norteiam o domínio ou disciplina; as necessidades do usuário; os tipos e características das entidades a

serem representadas e com isso mapear os requisitos funcionais do sistema para depois estabelecer os metadados.

Assim, o foco deste item está na análise conceitual e na determinação dos requisitos funcionais do sistema. Determinados pela modelagem conceitual, os requisitos funcionais auxiliarão o estabelecimento dos sistemas e padrões de metadados.

A análise conceitual é fruto do desenvolvimento de estudos que se iniciaram, a partir do final da década de 1970, sobre a modelagem e a construção de base de dados.

A modelagem conceitual, denominada também modelagem semântica, “[...] é um rótulo apropriado para a atividade geral de tentar representar significados” (DATE, 2004, p. 352), e a técnica mais utilizada refere-se à aplicação do método de entidade-relacionamento (E-R), desenvolvido e estudado pelos autores Codd, Peter Chen, Toby Teorey, James Fry, Ramez Elmasri, Coad, Yourdon, Shamkant Navathe, entre outros (DATE, 2004).

Sendo assim, a modelagem semântica é conhecida por muitas denominações, conforme aponta Date (2004): modelagem de dados, modelagem entidades/relacionamentos, modelagem de entidades, modelagem de objetos e ainda modelagem conceitual. Date (2004) aponta preferir a denominação modelagem semântica, pois as outras denominações apresentam-se conflitantes e muito específicas; por exemplo, o termo modelagem de dados é conflitante com o termo modelo de dados utilizado para indicar um sistema formal (como, por exemplo, o modelo relacional), porque tende a reforçar a idéia de modelo de dados utilizados apenas para estruturar dados; os termos modelagem de entidade-relacionamento e modelagem de entidades sugerem uma idéia específica para o problema (nesse sentido o autor denomina a abordagem do modelo entidade-relacionamento como uma técnica da modelagem semântica); e o termo modelagem de objetos parece estar em conflito com o termo entidade, já que significam a mesma coisa, além de o termo poder ser usado em outros contextos e com outros significados.

Date (2004) ressalta que a modelagem semântica incide sobre o desenvolvimento do projeto de banco de dados, e existem várias metodologias para o estabelecimento do modelo do banco de dados.

Nesta pesquisa o enfoque será dado ao modelo entidade-relacionamento, pois sua aplicação constitui-se em um modelo conceitual a ser refletido no sistema e que proporciona um conjunto de requisitos funcionais para o sistema de um determinado domínio.

Conforme já apontado, os metadados são elementos inerentes aos sistemas de bancos de dados e sua construção padronizada depende da aplicação de metadados coesos que, por

sua vez, necessitam de uma estruturação e determinação sólida nos sistemas de bancos de dados para serem processados com maior eficiência.

Dessa forma, destaca-se aqui a importância da modelagem conceitual no projeto lógico do banco de dados para que sejam mapeados e estabelecidos os requisitos funcionais do sistema e com ele os elementos (metadados) essenciais a serem implantados no banco de dados.

A modelagem conceitual é resultante de um processo de análise do sistema e determina o que o sistema deve conter, independente do modo como será implementado.

O modelo entidade-relacionamento é uma metodologia que mapeia as necessidades do sistema, por isso é importante que se conheçam o domínio e as necessidades do usuário para que se construa um modelo conceitual com requisitos sólidos. Dessa forma, ao final da aplicação do modelo entidade-relacionamento, há uma lista de requisitos funcionais do sistema em um determinado domínio (COAD; YOURDON, 1991).

Essa visão reflete uma mudança desenvolvida ao longo de décadas em relação ao estabelecimento dos sistemas de bancos de dados e mostra que, mais do que definir a programação a ser utilizada para a construção do banco de dados, deve-se pensar previamente no estabelecimento de requisitos para a construção desse sistema, ou seja, estabelecer e aplicar uma análise do domínio de problema, para que o sistema possa ser estabelecido com maior eficiência (COAD; YOURDON, 1991).

A consolidação do modelo conceitual se dá pela aplicação da metodologia do modelo entidade-relacionamento (E-R), gerando o mapeamento dos requisitos funcionais do sistema de um domínio.

De acordo com Coad e Yourdon (1991), o modelo entidade-relacionamento constitui-se como uma metodologia para capturar o conteúdo do domínio do problema e refletir na estrutura do banco de dados do sistema as necessidades ou os requisitos do mundo real. Sob outro ponto de vista, os autores destacam ainda que o modelo entidade-relacionamento pode ser considerado um modelo semântico de dados, pois mais do que estabelecer requisitos tecnológicos do sistema, o modelo E-R estabelece os requisitos semânticos de um domínio.

De acordo com o modelo entidade-relacionamento, existem três classes de objetos interligados: entidades, atributos e relacionamentos. De acordo com Chen (1990, p. 20), “uma entidade é uma ‘coisa’ que pode ser distintamente identificada”, e essa identificação é dada pelos atributos, que são as características de uma entidade. As associações entre as entidades são determinadas pelos relacionamentos e suas propriedades.

Silberschatz, Korth e Sudarshan (2006, p. 133-134) apontam que o início do projeto do banco de dados ocorre com a fase do projeto lógico, especificamente com a caracterização completa das “[...] necessidades de dados dos prováveis usuários do banco de dados”, e para isso destacam a necessidade de interação entre projetistas, especialistas de um domínio e usuários, a fim de que sejam determinadas as especificações das necessidades dos usuários. Após essa determinação, uma segunda etapa ainda do projeto lógico consiste na tradução dessas necessidades por meio da aplicação da metodologia conceitual do modelo entidade-relacionamento, dando origem a um modelo conceitual ou definição de requisitos funcionais. Esses processos constituem-se em um mapeamento conceitual para o modelo de dados a ser implementado no sistema. Após essa definição, vem a fase de implementação do projeto físico, aspecto que não se constitui objetivo desta pesquisa.

Para melhor entender o modelo entidade-relacionamento (E-R), apresentam-se brevemente a seguir algumas definições e considerações sobre o modelo.

- Entidades:

As entidades constituem-se como uma “coisa” ou “objeto” do mundo real que pode ser caracterizada, individualizada, distinguida de forma única e inequívoca (DATE, 2004; SILBERSCHATZ; KORTH; SUDARSHAN, 2006). Portanto, uma entidade possui um conjunto de características, propriedades, ou melhor, atributos e seus valores que a identificam de forma única.

As entidades podem ser concretas ou abstratas, e o conjunto de entidades do mesmo tipo compartilha os mesmos atributos; sua individualização é fornecida pelos valores dos atributos, ou seja, pela representação da informação.

- Atributos:

Os atributos são considerados por Silberschatz, Korth e Sudarshan (2006, p. 135) “propriedades descritivas” de um conjunto de entidades. Descrevem os valores, que por sua vez são as informações que representam a entidade.

O significado de atributo pode ser melhor entendido retomando-se a definição de metadados utilizada nesta pesquisa: **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.

Conforme pode ser percebido, os atributos são os metadados de um sistema de informação. Nesse sentido, uma importante questão a ser respondida é: quais atributos devem ser escolhidos para representar uma entidade? Peter Chen (1990, p. 27) responde afirmando que “[...] os atributos escolhidos devem ser capazes de identificar de forma absoluta as entidades”, isto é, devem-se selecionar todos os atributos necessários para a identificação e caracterização da entidade de forma única.

Além disso, é importante destacar que existe atualmente uma tendência na adoção de instrumentos para a padronização dos valores dos atributos, o que garante maior padronização e unicidade na representação. Essa tendência será abordada nos itens 5.1.2 e 5.2 e se constitui em um dos aspectos representacionais para a determinação dos metadados.

- Relacionamentos:

Os relacionamentos são as associações, ligações ou conexões estabelecidas por certas propriedades entre as entidades (CHEN, 1990). Os tipos de relacionamentos podem ser de um para um, um para muitos, de muitos para um e de muitos para muitos. Os relacionamentos auxiliam na determinação das associações entre as entidades e seus atributos, mapeando o domínio e estabelecendo relações entre requisitos.

Assim, considera-se fundamental o conhecimento do modelo E-R, pois acredita-se que ele se constitui como uma metodologia para representar conceitualmente um domínio. A partir da definição do modelo conceitual, é possível utilizar outros instrumentos para a construção padronizada de metadados, proporcionando uma determinação de metadados de forma mais consistente.

É importante a definição dos requisitos funcionais do sistema, pois eles refletem as necessidades informacionais do ponto de vista dos usuários finais, dos desenvolvedores do sistema e dos desenvolvedores de metadados, auxiliando na construção do sistema de banco de dados e também do esquema de metadados a ser utilizado (ZENG, QIN, 2004).

Em contraponto a essa importância, o estabelecimento de diferentes requisitos funcionais ocasiona o desenvolvimento de sistemas informacionais divergentes e, conseqüentemente, diferentes esquemas de metadados. Sua difusão possibilita o conhecimento da infraestrutura informacional do sistema, facilitando a construção padronizada não só do banco de dados, mas também dos esquemas de metadados de um determinado domínio (ZENG, QIN, 2004).

Neste sentido, destaca-se a importância do conhecimento sobre o modelo E-R e a relevância de uma determinação unificada dos requisitos funcionais do sistema, para que o modelo conceitual gerado seja utilizado de forma unificada em um domínio.

Se houver um mapeamento incorreto ou estabelecimento dos requisitos de forma equivocada, e um valor for colocado em um elemento de metadado que não o representa, ou seja, em um atributo que não representa o significado absoluto daquele determinado valor, a informação não será recuperada adequadamente.

É importante que o projetista de banco de dados conheça não só o domínio de aplicação e as necessidades dos usuários, mas principalmente os princípios que norteiam este domínio. Isso criará um conjunto de requisitos funcionais consistentes em um domínio.

Após o estabelecimento dos requisitos funcionais do sistema, é preciso determinar quais atributos são essenciais para o esquema de metadados. A identificação dos elementos essenciais corresponde a uma lista de atributos que serão utilizados para compor o esquema de metadados. Essa lista de elementos contempla as necessidades dos usuários, as características da coleção e os requisitos funcionais do sistema, de acordo com as opções disponíveis e identificadas nos requisitos funcionais. Nesse sentido, diferentes domínios apresentam elementos divergentes (ZENG, QIN, 2004). Ressalta-se aqui que é importante haver requisitos funcionais estabelecidos para cada domínio, constituindo-se assim em modelos conceituais padrões para os sistemas.

É fundamental que se estabeleçam esses atributos ou elementos essenciais⁵ com as seguintes características, de acordo com Zeng e Qin (2004, p. 93):

- Elementos essenciais
 - Explicação e descrição do elemento
 - Exemplo
 - Implementação (M - obrigatório, O - opcional, R - repetível)
- Uma coleção de elementos essenciais conduz para o próximo passo de um projeto de metadados, que será definir o real esquema de metadados.

A próxima etapa consiste no estabelecimento do esquema de metadados, por meio da definição do conjunto de elementos reais e essenciais para a representação de uma entidade (ZENG, QIN, 2004).

⁵ Os autores Zeng e Qin (2008) denominam os elementos essenciais como “*desired elements*”, que poderiam ser entendidos como os elementos que se deseja ter no esquema. Assim, para esta pesquisa o termo “*desired elements*” foi traduzido como elementos essenciais, e se constituem nos elementos essencialmente necessários para representar uma entidade de forma única.

5.1.2 ESQUEMAS DE METADADOS: conjunto de elementos e espaço do valor

Uma representação de qualidade depende da aplicação de um esquema de metadados consistente, que por sua vez depende da determinação dos metadados. Assim, um esquema de metadados pode ser entendido como o conjunto de elementos descritivos (atributos codificados ou metadados) e regras de descrição que compõem um padrão de metadados (MÉNDEZ RODRÍGUES, 2002).

Existem três opções para a utilização de um esquema de metadados: a) criar um novo esquema, b) adotar um esquema já existente, ou c) modificar um esquema já existente. O processo de determinação do esquema, em qualquer uma dessas opções, exige constantes revisões e adaptações, pois deve estar em conformidade com requisitos funcionais do sistema e princípios de um domínio para que possam determinar metadados de modo coeso (ZENG, QIN, 2008).

Embora existam diferentes tipos de esquemas de metadados, há também certa semelhança entre eles, relacionada aos aspectos de estrutura, semântica e sintaxe, que conferem ao esquema uma formalidade e a possibilidade de representação padronizada. Assim, abordam-se neste item as características consideradas mais relevantes em relação aos aspectos de estrutura, sintaxe e semântica, para a consolidação dos metadados e do esquema de metadados.

Contudo, é preciso entender inicialmente o que são esquemas e sua importância para a determinação dos metadados e esquemas de metadados. Para isso, apresenta-se o significado e a explicação dos seguintes termos: *Schema* e *Scheme*. Os dois termos são traduzidos para o português como esquema, porém, apresentam significados distintos.

De acordo com Méndez Rodrigues (2002, p. 349-350, tradução nossa), o termo “*Schema*” possui alguns significados que variam de acordo com o âmbito de sua utilização: no âmbito dos metadados e das bases de dados, é um termo utilizado “[...] para descrever a estrutura de uma tabela em uma base de dados relacional.” No âmbito da linguagem em XML e do modelo RDF, o termo *schema* é utilizado para definir estruturas e vocabulários: “Em XML, um *schema* é um modelo para descrever a estrutura da informação de toda uma classe de documentos.” Já no modelo ou arquitetura RDF “[...] um *schema* define um mecanismo para declarar as propriedades dos elementos, assim como as relações entre as propriedades e os recursos, e também a semântica para realizar tudo isso.” Em um sentido mais amplo e sobretudo ligado aos metadados, o termo *schema* refere-se “[...] ao conjunto de elementos e

das regras que constituem um modelo de metadados”, ou seja, à definição de atributos e regras que se constituem no esquema de metadados (conjunto de atributos codificados – metadados) do padrão de metadados.

Já o termo “*Scheme*” apresenta um significado complexo; está relacionado com um ponto de vista técnico da linguagem de marcação, mas apresenta um significado que vai além de um componente estrutural da linguagem. Para Méndez Rodrigues (2002, p. 350, tradução nossa), *Scheme* refere-se “[...] ao esquema que se deve utilizar para interpretar o valor de uma propriedade” ou atributo. Em outras palavras, refere-se à especificação dos valores nos metadados: ordem, disposição lógica e um conjunto de valores possíveis determinados por padrões externos ao esquema de metadados e o modo como isso é codificado em uma linguagem legível por máquina.

Os autores Zeng e Qin (2008, p. 131) definem que o termo *schema* na comunidade de metadados

[...] significa a definição semântica e estrutural dos elementos de metadados e as relações entre esses elementos. Essas representações definem qual o significado de um elemento de metadados e como será codificado em um formato legível por máquina. [...]

Deste modo, identifica-se nesta citação a presença dos dois tipos de esquemas. O esquema do tipo *schema* define os atributos e as regras sob os aspectos semânticos e estruturais. Já o esquema do tipo *scheme* define a sintaxe de codificação dos dados, o que auxilia no estabelecimento da estrutura e da semântica (significado) dos atributos e valores.

Dessa forma entende-se que um esquema de metadados (*schema*) é composto por outros tipos de esquemas (os *schemes*) que vão estruturar a descrição por meio da ordem e disposição lógica das declarações (*statements*) que compõem o conjunto de elementos (*element set*) do esquema de metadados do padrão.

Neste cenário, cada tipo de esquema (*schema* e *scheme*) apresenta uma finalidade e importância, sendo complementares entre si, pois ajudam no estabelecimento de metadados e esquemas de metadados consistentes.

É preciso entender a estrutura e os componentes básicos dos esquemas de metadados, pois por meio dele é possível compreender a determinação dos metadados em relação aos aspectos tecnológicos.

Sendo assim, destacam-se os seguintes componentes do esquema de metadados: o conjunto de metadados (*element set* – ou metadados prescritos) e o espaço do valor (*value spaces*). Conforme se verifica a seguir, cada um desses componentes irá apresentar estruturas

e esquemas próprios, que auxiliarão na determinação dos aspectos estruturais, semânticos e de sintaxe para a determinação dos metadados e esquemas de metadados do padrão.

5.1.2.1. CONJUNTO DE METADADOS (*ELEMENT SET*)

Para compreender o conjunto de elementos, é necessário primeiramente definir o termo esquema de metadados (*metadata schema*). De acordo com Zeng e Qin (2008, p. 323, tradução nossa), esquema de metadados é

Uma especificação legível por máquina que define a estrutura, codificação sintática, regras, e formatos para o conjunto de elementos de metadados em um esquema de linguagem formal. Na literatura o termo “esquema de metadados” refere-se comumente ao conjunto completo de elementos de metadados como também a codificação dos elementos e estrutura em uma linguagem de marcação.

Explicando essa definição pode-se dizer que o esquema de metadados (*metadata schema*) é composto pelo conjunto de elementos (*element set*), que se constitui em uma especificação, ou uma descrição detalhada, ou ainda um conjunto de elementos prescritos que descrevem um recurso ou entidade.

Os elementos ou os metadados que descrevem uma entidade e são especificados e individualizados por meio de declarações (*statements*) apresentam uma estrutura em forma de par propriedade-valor (ou atributo-valor). A estrutura da declaração irá compor a estrutura de descrição do esquema de metadados e juntamente com os aspectos sintáticos de codificação dos dados vão garantir que os aspectos semânticos dos dados sejam refletidos nas declarações, contribuindo para o estabelecimento de uma estrutura formal⁶ de descrição.

Existem dois componentes básicos no conjunto de elementos do esquema: “[...] Semântica (definições dos significados dos elementos e suas relações) e Conteúdo (declarações ou instruções do que e como os valores devem ser atribuídos aos elementos)” (ZENG, QIN, 2008, p. 93).

Um atributo representa uma propriedade ou característica de uma entidade. Para ele existir em um esquema, deve ser codificado e estruturado, compondo um elemento de metadados. Individualmente o elemento é formado pela codificação do atributo, representada

⁶ Entende-se por estrutura formal a codificação lógica e coerente dos metadados em um esquema de metadados.

por uma etiqueta (*label*), nome do atributo e sua definição (ZENG, QIN, 2008). A seguir apresenta-se uma breve explicação dos componentes básicos do conjunto de elementos:

- Definição: é uma declaração que representa o conceito essencial e a natureza do atributo. A definição deve conter apenas a informação necessária para defini-lo, por isso deve ser concisa e clara, e se possível declarada em uma oração breve. “A definição declara o que o conceito é, no lugar do que não é.” (ZENG, QIN, 2008, p. 94).
- Estrutura: refere-se à organização dos elementos que geralmente são estruturados de duas formas: a) estrutura plana - não há indicação de subelementos, mas podem ser adicionados posteriormente na forma de qualificadores, que refinam, restringem e especificam a semântica do elemento; b) estrutura hierárquica - apresenta metadados multivalorados, ou seja, elementos divididos em subelementos, que por sua vez podem ser especificados em outros subelementos (ex.: atributo data e os vários tipos de datas, data de publicação, reimpressão etc) (ZENG, QIN, 2008).
- Restrição de conteúdos: refere-se a padrões que refinam e especificam as descrições no intuito de ampliar o significado dos dados. O termo restrição de conteúdo refere-se à especificação dos dados, no sentido de proporcionar uma descrição mais detalhada. As restrições de conteúdos são definidas por regras e padrões, e um exemplo a ser citado são as regras para restrição dos conteúdos dos valores dos elementos (ZENG, QIN, 2008).
- Apresentação: a apresentação do conjunto de elementos em uma especificação pode variar de acordo com a linguagem de marcação utilizada, porém essa é uma versão utilizada para o processamento legível por máquina. “O formato mais compreensível é obtido com a apresentação dos elementos básicos sem especificação de codificação, de forma que a estrutura, elementos, definições, e conteúdos são expressos de modo coerente e claro” (ZENG, QIN, 2008, p. 96).
- Codificação / ligação: os elementos considerados essenciais são codificados para compor o conjunto de elementos do esquema. A codificação define a estrutura e a sintaxe dos elementos em uma linguagem formal para ser legível por máquina. Mas é possível que o mesmo conjunto de elementos seja codificado e ligado com padrões diferentes de codificação, facilitando assim o intercâmbio de dados (ZENG, QIN, 2008).

Deste modo, entende-se que os aspectos estruturais, semânticos e sintáticos estão presentes em um esquema de metadados da seguinte forma: a estrutura determina a disposição dos atributos que representam uma entidade no conjunto de elementos do padrão, e na

estrutura estão embutidos os relacionamentos entre os elementos, subelementos e entre as entidades; a sintaxe possibilita a codificação dos elementos determinados na estrutura de descrição, estabelece também a ordem lógica dos valores, além da codificação dos elementos em um formato legível por máquina. Tanto a estrutura como a sintaxe expressam os aspectos semânticos: os aspectos semânticos da estrutura estão relacionados à definição do significado atributo e às conexões entre os elementos, subelementos e entidades; e os aspectos semânticos referentes à sintaxe estão relacionados à ordem lógica de codificação e expressão dos valores nos elementos de metadados (ZENG, QIN, 2008).

Entretanto, a expressão dos valores constitui-se como algo complexo, pois exige em alguns momentos uma padronização exterior à estrutura do esquema de metadados. A sintaxe estabelece a lógica para expressar valores em um elemento de descrição, mas como atribuir um valor semântico ao espaço do valor? Quais os valores admissíveis em cada espaço? (ZENG, QIN, 2008).

Essas questões são respondidas por regras definidas em esquemas externos de representação. Sendo assim, aborda-se a seguir a questão do espaço do valor nos esquemas de metadados.

5.1.2.2 ESPAÇO DO VALOR (*VALUE SPACES*)

O espaço do valor constitui-se como um componente importante; nele será expressa a representação de uma entidade dada pelo atributo ou metadado. Dessa forma, a padronização do espaço do valor torna-se essencial para proporcionar uma representação de qualidade.

De acordo com Zeng e Qin (2008, p. 103) o espaço do valor pode ser definido como “[...] um conjunto de valores para um determinado tipo de dados. [...] Em padrões de metadados, o espaço do valor expressa o conjunto de valores e/ou regras de especificação para cada elemento e posição no conjunto de elementos.” (ZENG, QIN, 2008, p. 103).

Em outras palavras, a definição do espaço do valor ocorre por meio de uma sintaxe para expressar os valores nos elementos e por meio de esquemas de codificação que definem regras e atribuem valores. Esses esquemas unificam o espaço do valor e devem ser utilizados, em alguns casos, juntamente com o esquema de metadados.

O estabelecimento e o uso de padrões para o espaço do valor refletem a tendência atual de padronização dos esquemas de metadados, pois não basta apenas estabelecer e

utilizar um esquema de metadados, é necessário utilizar também esquemas para padronização valores a fim de garantir uma representação consistente.

Conforme observado, a determinação padronizada dos valores depende de esquemas externos. Inicia-se aqui a convergência de um outro aspecto para a determinação dos metadados, os aspectos representacionais.

Antes de abordá-los, é necessário apontar que o estabelecimento do esquema de metadados deve estar diretamente relacionado com a aplicação correta dos princípios. Entretanto, determinar o esquema é uma tarefa complexa, principalmente em relação às especificidades de um domínio, pois torna-se difícil manter a simplicidade na estrutura do esquema quando o domínio exige um nível de detalhe ou especificidade.

É fundamental ressaltar a necessidade de conhecimento dos princípios que norteiam o domínio e a importância do estabelecimento dos aspectos tecnológicos para que possam ser criados esquemas de metadados coerentes, proporcionando assim representações adequadas dos recursos informacionais.

5.2 ASPECTOS REPRESENTACIONAIS PARA A DETERMINAÇÃO DOS METADADOS: esquemas de codificação

O aspecto representacional constitui-se como um fator significativo na determinação e padronização dos metadados. Conforme destaca Méndez Rodríguez (2002), esses aspectos referem-se aos instrumentos de representação na padronização dos dados legíveis por humanos.

Nos esquemas de metadados, o aspecto representacional refere-se ao uso de esquemas de codificação nos elementos. Esses esquemas são representados por padrões indicados na estrutura do esquema para estabelecer os valores nos espaços do valor, normalizando os aspectos semânticos e sintáticos da representação na estrutura do esquema de metadados. Desse modo, os esquemas de codificação determinam os aspectos representacionais por meio de outros padrões externos ao esquema de metadados, porém associados a ele.

Os esquemas de codificação representam um importante papel não só na padronização dos valores, mas também na determinação dos metadados, mostrando que aspectos tecnológicos e representacionais estabelecem uma relação intrínseca, conforme apontado anteriormente.

De acordo com Zeng e Qin (2008), podemos encontrar os seguintes tipos de padrões utilizados como esquemas de codificação:

- a) Padrões de estrutura de dados (*data structure standards*): corresponde ao conjunto de elementos do esquema de metadados, explicado anteriormente nesta pesquisa;
- b) Padrões de conteúdo de dados (*data content standards*): refere-se a “padrões de conteúdos de dados [que] guiam a escolha dos termos e definem a ordem, sintaxe e forma na qual os valores de dados devem ser apresentados na estrutura de dados” (ZENG, QIN, 2008, p. 127; MÉNDEZ RODRÍGUEZ, 2002). De acordo com Baca (1998, p. 37), esses padrões são compostos por “[...] regras e diretrizes que regem a ordem, a sintaxe e a forma em que se introduzem os valores dos dados dentro de uma estrutura de dados. [...]”. Alguns padrões de conteúdo de dados indicados no domínio Web foram especificamente desenvolvidos no domínio bibliográfico.
- c) Padrões de valores de dados (*data value standards*): constituem-se como “um padrão projetado e usado para controlar ou restringir valores na declaração de metadados gerando indexação e recuperação” (ZENG, QIN, 2008, p. 320). Como exemplos desses padrões citam-se: tesouros, listas de autoridade, listas controladas do esquema, léxicos, esquemas de classificação, vocabulários controlados, listas de títulos, entre outros (BACA, 1998; ZENG, QIN, 2008). Esses padrões determinam os valores possíveis a serem admitidos em um espaço do valor, auxiliando, assim, no estabelecimento de relações semânticas entre atributos e entidades, possibilitando a construção de representações consistentes.
- d) Padrões de intercâmbio de dados (*data exchange standards*): “uma especificação para armazenamento, acesso e transmissão de dados referentes a “formatos” quando discutidos no contexto de intercâmbio de dados.” (ZENG, QIN, 2008, p. 320). Os padrões para intercâmbio de dados, também denominados formatos, apresentam uma conotação diferente dos termos utilizados nesta pesquisa. Esses padrões referem-se aos formatos de codificação da informação utilizados no contexto de intercâmbio de dados e podem ser ou não projetados separadamente ao esquema do padrão. No padrão MARC 21, por exemplo, a estrutura de intercâmbio do registro é dada pela ANSI/NISO Z39.50 e pela ISO 2709, que determina a forma legível por máquina dos dados bibliográficos a serem intercambiados no sistema. Para os padrões de metadados do domínio Web, é indicada a utilização da linguagem de marcação XML, por ser uma linguagem flexível para intercâmbio de dados, juntamente com a arquitetura RDF (Resource Description Framework) (ALVES, 2005; ZENG, QIN, 2008). Embora os padrões de intercâmbio de dados estejam mais ligados aos aspectos tecnológicos de intercâmbio de dados, são

destacados aqui por necessitarem da determinação dos aspectos representacionais para que o intercâmbio de dados possa ocorrer de modo consistente.

Existem várias denominações na literatura para designar os tipos de esquemas de codificação. Nesta pesquisa, entende-se a relação entre esquemas de codificação e os padrões utilizados para a codificação conforme apresentado no quadro a seguir:

ESQUEMAS DE CODIFICAÇÃO	PADRÕES	DOMÍNIO WEB PADRÃO DC	DOMÍNIO BIBLIOGRÁFICO PADRÃO MARC21
ESQUEMAS DE CODIFICAÇÃO DE CONTEÚDO	Padrões de estrutura de dados (<i>data structure standards</i>)	Esquema de metadados: estrutura do conjunto de elementos representada pelas declarações (<i>statements</i>)	Esquema de metadados: estrutura descritiva em diretórios, campos, subcampos, código de subcampo, indicadores
	Padrões de intercâmbio de dados (<i>data exchange standards</i>)	Esquema da Linguagem XML Protocolo (OAI-PMH)	Protocolo ANSI/NISO Z39.50 Norma ISO 2709
ESQUEMAS DE CODIFICAÇÃO DE VALORES	Padrões de conteúdo de dados (<i>data content standards</i>)	Fornecido no próprio esquema de metadados do padrão DC, pelas diretrizes de boas práticas ou pelo uso de padrões externos como os existentes no domínio bibliográfico	Padrões ou códigos elaborados para o TDI. Ex.: AACR, ISBDs
	Padrões de valores de dados (<i>data value standards</i>)	Listas controladas do esquema, como por exemplo: DCMI Type Vocabulary, determinado como esquema de codificação do elemento Type no esquema DC, Folksonomia (DCMI Social Tagging Community)	Tesauros, listas de autoridade, léxicos, esquemas de classificação, vocabulários controlados

QUADRO 6: Esquemas de codificação e padrões correspondentes.

Fonte: do autor.

O quadro 6 demonstra que os esquemas de codificação são subdivididos em esquemas para codificação de conteúdo e esquemas para codificação de valores.

Os esquemas de codificação de conteúdo são constituídos pelos padrões de estrutura dos dados e padrões de intercâmbio de dados. Os padrões de estrutura de dados correspondem ao esquema de metadados de um padrão, o qual apresenta um conjunto de elementos prescritos, sua definição e um comentário do que deve ser apresentado. Essa estrutura é representada pelas declarações (*statements*). Os padrões de intercâmbio de dados determinam a codificação dos dados nos processos de intercâmbio, ou seja, como são lidos e “compreendidos” por máquina. Para isso são determinados por outros esquemas ou padrões,

tais como o esquema da linguagem de marcação XML, o protocolo OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting – protocolo de coleta de metadados em arquivos abertos), padrões como o protocolo ANSI/NISO Z39.50 e a norma ISO 2709 para intercâmbio de dados bibliográficos.

Os esquemas de codificação de valores são constituídos pelos padrões de conteúdo de dados e padrões de valores de dados. Os padrões de conteúdo de dados correspondem aos esquemas ou padrões externos que determinam regras e diretrizes que estabelecem a ordem, a sintaxe e a forma como os dados vão ser inseridos no esquema. Os padrões para valores de dados determinam a codificação dos valores nos espaços do valor do esquema, e são utilizados para restringir (no sentido de delimitar e especificar) os possíveis valores admitidos nos esquemas de metadados.

De modo geral, pode-se considerar que os esquemas de codificação estabelecem regras para a descrição do conteúdo nos espaço do valor, regras de codificação e sintaxe dos dados e as possíveis formas ou valores permitidos aos valores nos espaços de valor, levando em consideração a definição do atributo e a questão do intercâmbio dos metadados. Alguns desses esquemas e padrões são específicos de um domínio, como o domínio bibliográfico, mas estão sendo indicados em padrões de metadados para Web, como o padrão DC, por também abordarem a descrição de recursos informacionais.

Considera-se também que os metadados estruturados e, principalmente, os metadados ricos destacados nesta pesquisa apresentam os aspectos representacionais consolidados, e a utilização de esquemas ocorre em maior ou menor grau. A diferença entre esses padrões é que no caso dos padrões estruturados, em especial o padrão DC, os esquemas de codificação são opcionais em alguns elementos, enquanto em padrões ricos como o MARC, o uso de esquemas de codificação é uma exigência do domínio, pois garante a especificidade e exatidão na representação.

Esses esquemas devem ser aceitos no domínio ao qual pertencem e ser utilizados como instrumentos que determinam regras e diretrizes para promover uma representação e armazenamento adequado e consistente dos dados, de modo que possam facilitar a busca, recuperação, intercâmbio dos dados e interoperabilidade entre os sistemas (MÉNDEZ RODRÍGUEZ, 2002).

Um padrão de metadados pode apresentar um bom esquema de descrição, mas pode proporcionar uma representação insuficiente se não forem utilizados esquemas de codificação para unificação dos dados e valores, ou seja, esquemas que normalizem os aspectos representacionais dos metadados, os quais garantem que a semântica dos dados será

devidamente representada no esquema de metadados, possibilitando o intercâmbio mais consistente de dados e uma recuperação mais eficiente.

Neste contexto, os guias de boas práticas e os princípios a serem seguidos em um domínio também se destacam como instrumentos importantes no estabelecimento de metadados e seus valores, pois determinam as diretrizes a serem seguidas para a construção de representações (ZENG, QIN, 2008).

Méndez Rodríguez (2002) considera que em contextos genéricos, como ocorre no domínio Web, a padronização dos metadados ocorre com a aplicação consistente de MetaTag(s), entretanto, não engloba somente as ferramentas de criação de etiquetas. De acordo com a autora, “Os metadados requerem uma consistência em seu conteúdo e estilo para assegurar o intercâmbio de informação” (MÉNDEZ RODRÍGUEZ, 2002, p. 195, tradução nossa). Isso reforça a necessidade do uso de esquemas de codificação que determinam questões representacionais e garantem uma construção e determinação coerente de metadados.

Nesse sentido, a determinação padronizada dos metadados no domínio Web consiste no estabelecimento dos seguintes itens: modelagem conceitual do sistema, determinação dos requisitos funcionais, mapeamento dos elementos essenciais, determinação da estrutura de descrição do esquema, determinação dos espaços do valor e determinação dos esquemas de codificação para padronização dos espaços do valor. Apoiando esses itens existem ainda outras ferramentas tecnológicas como, por exemplo, os padrões para intercâmbio de dados e as linguagens de marcação que conferem ao esquema do padrão a flexibilidade necessária ao intercâmbio dos dados.

No domínio bibliográfico representado pelo padrão de metadados MARC 21, o estabelecimento dos metadados e do esquema ocorre de modo mais complexo, por exigir maior formalidade e especificidade na representação dos recursos. A especificidade e o detalhamento na descrição são características do domínio bibliográfico, portanto os instrumentos que padronizam as representações no padrão MARC 21 (esquemas de codificação) são fundamentais para proporcionar uma representação consistente no esquema de metadados do padrão, facilitando assim o intercâmbio de dados. Embora isso já esteja consolidado, é necessário o conhecimento desses instrumentos para que possam ser criadas posteriormente ferramentas que garantam maior interoperabilidade entre sistemas.

Podemos considerar que a determinação padronizada dos metadados e do esquema de metadados do padrão MARC 21 ocorreu antes mesmo de sua existência de forma automatizada. Os aspectos representacionais do padrão MARC 21: metadados essenciais, o

conjunto de elementos do esquema (elementos prescritos), os espaços do valor, a sintaxe de codificação dos dados, os instrumentos de representação de valores (esquemas de codificação) já haviam sido determinados previamente em instrumentos destinados ao tratamento descritivo da informação desenvolvidos ao longo do tempo na disciplina de Catalogação.

O estabelecimento de princípios de Catalogação, dos instrumentos de representação (códigos) e a determinação das regras foram instrumentos essenciais na criação padronizada de metadados no domínio bibliográfico. Com o passar do tempo, o tratamento descritivo da informação passou de manual para automatizado, e alguns desses padrões e instrumentos são refletidos no meio tecnológico pelo esquema do padrão MARC 21 e outros são utilizados como instrumentos de padronização das representações na estrutura do esquema (esquemas de codificação).

Analisando essas considerações e as características dos padrões de metadados estudados nesta pesquisa, considera-se que os padrões de metadados são compostos pelo esquema de descrição (conjunto de elementos prescritos) e pelos esquemas de codificação associados, que se constituem em padrões, diretrizes e regras que auxiliam no estabelecimento dos metadados e seus valores em relação aos aspectos representacionais da informação em um domínio. Acredita-se que um padrão de metadados engloba em sua estrutura descritiva aspectos tecnológicos e representacionais, compondo uma estrutura completa para descrição e intercâmbio de dados, possibilitando o tratamento descritivo da informação e a interoperabilidade entre sistemas de forma adequada.

Conforme observado neste capítulo, os aspectos tecnológicos e representacionais estão intimamente relacionados no intuito de proporcionar uma representação coerente e concisa dos recursos informacionais.

Os aspectos relacionados com os padrões de metadados estruturados e os padrões de metadados ricos, bem como os aspectos tecnológicos e representacionais destacados neste capítulo apontam os requisitos necessários para a determinação de metadados e as características dos instrumentos de descrição (padrões de metadados) em cada domínio destacado nesta pesquisa.

A representação informacional no domínio Web congrega todos os princípios, ferramentas tecnológicas e aspectos representacionais emergentes em relação aos metadados.

Em domínios específicos como o domínio bibliográfico, por exemplo, a representação informacional depende dos princípios e instrumentos estabelecidos no domínio e sua interação com as características emergentes dos metadados. Essa interação entre o domínio bibliográfico e os aspectos emergentes de representação do domínio Web pode ser observada

na versão do padrão MARC 21 em XML. Isso demonstra que, mesmo se tratando de um domínio específico, com características específicas de descrição, é necessária a integração com o domínio Web para a construção de estruturas mais flexíveis de representação e que sejam correspondentes às atuais necessidades dos sistemas de informação.

Sendo assim, observa-se que o domínio bibliográfico vem passando por intensas alterações na tentativa de adequar suas características de descrição às novas necessidades informacionais impostas pelos sistemas de informação devido à criação de metadados no domínio Web.

Neste cenário, presencia-se uma reavaliação das teorias, dos princípios, dos métodos e instrumentos de representação no domínio bibliográfico, com o intuito de proporcionar uma representação mais consistente, tendo em vista os novos ambientes informacionais digitais.

Sabe-se que as representações vão depender do tipo de entidade que se busca representar, pelo tipo de necessidade a que se busca atender, pelo tipo de ambiente informacional e principalmente pelo domínio ao qual pertence (ORTEGA, 2009).

Tendo em vista as mudanças que estão ocorrendo no domínio bibliográfico, quais as teorias, os princípios, os fundamentos e os instrumentos que norteiam a construção de representações? Quais requisitos devem ser seguidos para a determinação dos metadados e a construção consistente de representações no domínio bibliográfico?

Apresenta-se no capítulo a seguir uma análise das teorias, dos princípios, dos fundamentos e instrumentos utilizados no TDI.

6 TEORIAS, PRINCÍPIOS, FUNDAMENTOS, MÉTODOS E INSTRUMENTOS DO TRATAMENTO DESCRITIVO DA INFORMAÇÃO NO DOMÍNIO BIBLIOGRÁFICO

Ao longo dos séculos, muitos indivíduos e organizações contribuíram para o desenvolvimento do tratamento descritivo da informação no domínio bibliográfico, com o intuito de estabelecer um modo padronizado para o processo de representação.

Assim, a busca de uma consolidação consistente para o TDI possibilitou o desenvolvimento de princípios e fundamentos da disciplina, dos métodos para o processo de TDI e dos instrumentos utilizados na construção de representações.

Neste cenário, as tecnologias de informática interagem estrategicamente com as teorias, com os princípios, com os fundamentos, com os métodos e com os instrumentos, ora como facilitadora do processo, ora como agente principal de modificação do TDI, ora como norteadora de princípios, instrumentos e sistemas. Por isso, é importante contextualizar os principais conceitos que serão estudados neste capítulo.

Para esta pesquisa, considera-se teoria o “conjunto de princípios fundamentais duma arte ou duma ciência”, “opiniões sistematizadas” (FERREIRA, 2004). Em relação à teoria de Catalogação, considera-se teoria o conjunto de conhecimentos sistematizados a partir da prática da catalogação e estabelecidos por meio dos princípios fundamentais da disciplina. (FERREIRA, 2004).

Os princípios são considerados a “causa primária”, “preceito, regra, lei”, “origem de algo, de uma ação ou de um conhecimento”. Os princípios abordados aqui se referem aos princípios e preceitos (regras de proceder) que determinam o domínio bibliográfico. (FERREIRA, 2004).

O fundamento é um termo utilizado muitas vezes como sinônimo de princípio (no sentido de origem), refere-se ao conjunto de enunciados (princípios). O fundamento é razão de tal enunciado no sentido de ser sua explicação racional (MORA, 2000). Podemos entender também como “o conjunto dos princípios básicos de um ramo de conhecimento [...]” (FERREIRA, 2004).

Quanto ao método, o termo utilizado refere-se ao modo de proceder ou aos procedimentos do processo de TDI.

E os instrumentos são definidos como “recurso empregado para se alcançar um objetivo, conseguir um resultado; meio” (FERREIRA, 2004) e podem ser entendidos como os instrumentos, as diretrizes, os códigos, os padrões de metadados e os esquemas de codificação associados que possibilitam a representação do recurso no TDI.

Atuando como agentes de mudança no TDI, as tecnologias de informática vêm contribuindo para uma reavaliação da teoria, dos princípios, dos fundamentos, dos métodos e dos instrumentos de representação do domínio bibliográfico. Essa reavaliação vem sendo desenvolvida com o objetivo de tornar mais consistente o processo de TDI em relação aos novos ambientes informacionais digitais e às novas necessidades dos usuários diante das tecnologias.

A padronização de representações sempre foi um objetivo do TDI, fato também observado no domínio Web, com a busca de instrumentos que padronizem os metadados. Os aspectos tecnológicos e representacionais tratados no capítulo anterior reforçam a constatação desse fato: por meio de esquemas que representam os espaços do valor e esquemas que auxiliam a determinação conceitual dos metadados, busca-se estabelecer uma consistência na representação nos padrões de metadados estruturados.

Assim, no domínio bibliográfico essa tendência de padronização sempre esteve presente no intuito de estabelecer estruturas coerentes de descrição para proporcionar representações mais consistentes. Entretanto, com as atuais mudanças causadas pelas tecnologias de informática, o domínio bibliográfico vem passando novamente por uma reavaliação e aprimoramento.

Embora o TDI seja um processo prático, acredita-se nesta pesquisa que ele não se consolida sem uma teoria previamente estabelecida e subjacente à prática de catalogação desenvolvida ao longo do tempo. Assim, os estudos teóricos sobre a prática de catalogação consolidaram princípios, fundamentos, métodos e instrumentos que possibilitam um tratamento descritivo da informação de modo padronizado e consistente.

O intuito deste capítulo é demonstrar que os métodos atuais de TDI, aliados às mudanças proporcionadas pelas tecnologias de informática, são adequados para a construção de representações padronizadas no domínio bibliográfico, sem a necessidade do uso de outros padrões que não sejam os já determinados no domínio.

Considera-se que o TDI apresenta o referencial teórico e prático que orienta a definição dos metadados no domínio bibliográfico. Assim, com base nos aspectos estudados nesta pesquisa sobre os metadados e os aspectos a serem tratados aqui sobre o TDI, propõe-se uma categorização dos instrumentos utilizados na determinação de metadados e na construção de representações no domínio bibliográfico.

6.1 CONTRIBUIÇÕES PARA UMA TEORIA DA CATALOGAÇÃO

Muitos conceitos e princípios foram estabelecidos na história dos catálogos e da Catalogação, em decorrência de várias mudanças no cenário científico tecnológico e da constante necessidade de se estabelecer um procedimento padronizado para o TDI. A busca da consolidação eficiente de um tratamento descritivo da informação possibilitou a concretização de uma teoria que norteasse a construção de princípios e fundamentos para a Catalogação. Essa teoria pode ser observada pela história dos catálogos e da Catalogação, e contribuiu para a determinação de princípios, fundamentos, regras e instrumentos de representação, sendo que os mais influentes e importantes perduram até os dias de hoje, conforme poderá ser observado neste capítulo.

Assim, destacam-se a seguir algumas idéias e contribuições estabelecidas pelos principais teóricos na Catalogação, e os destacados aqui foram escolhidos com base nos estudos de Chan (2007), Barbosa (1978), Mey e Silveira (2009). Após sua contribuição, apresenta-se uma breve análise.

6.1.1 PRINCÍPIOS DE PANIZZI

Anthony Panizzi elaborou o primeiro código de catalogação conhecido como “As 91 regras de Panizzi”. Nelas estavam subentendidas as seguintes ideias e princípios: o objetivo do catálogo de biblioteca, a normalização da entrada principal (pessoa e entidade), a condição de obras modificadas ou adaptadas para diferenciar as originais, exigências para uniformidade na aplicação das regras de catalogação, entre outros (CHAN, 2007).

Análise: embora tenham ocorrido muitas alterações desde o surgimento do primeiro código, podem-se considerar as regras de Panizzi como precursoras das regras de catalogação atuais. Portanto, sua maior contribuição refere-se ao início do estabelecimento de padronização por meio de regras de catalogação; a quantidade de regras, questionada por muitos estudiosos, reflete a necessidade de especificidade que o domínio bibliográfico exige.

6.1.2 PRINCÍPIOS DE JEWETT

Charles C. Jewett elaborou, com base nas regras de Panizzi, um código de catalogação para Smithsonian Institution nos Estados Unidos.

De acordo com Barbosa (1978, p. 28, grifo do autor) a importância de seu nome está ligada “[...] à elaboração de um *catálogo coletivo*, onde pretendia reunir as informações sobre coleções existentes nas bibliotecas americanas e imprimi-las pelo processo de estereotipia.”

De acordo com Chan (2007, p. 48) pode-se considerar Jewett um visionário do “catálogo universal” ou catálogo coletivo. Ele idealizou duas linhas de ação, que podem ser **pensadas** como dois princípios que se estabeleceram posteriormente na catalogação:

O primeiro era o uso de modelos de entradas e separadas em placas de estereotipia para facilitar a produção do catálogo e, além disso, reduziria drasticamente o custo com a produção e manutenção do catálogo, particularmente com a preparação de novas edições do catálogo em livro que era a forma de catálogo predominante naquele tempo. A segunda proposta de Jewett era a catalogação cooperativa, um desenvolvimento visto como necessário para construir um catálogo universal. [...] (CHAN, 2007, p. 48).

Análise: conforme se percebe, esses ideais visionados por Jewett se refletiram, mais tarde, no catálogo universal ou catálogo coletivo e na idéia de catalogação cooperativa, ideal que só se consolida efetivamente com a inserção das tecnologias de informática nos processos de catalogação.

6.1.3 PRINCÍPIOS DE CUTTER

Charles Ami Cutter foi um teórico de significativo destaque no campo da Catalogação; elaborou as “Rules for a dictionary catalog” (regras para um catálogo dicionário) no qual estabelecia não só regras para “[...] entradas por autor e título, para a parte descritiva, cabeçalhos de assunto, alfabetação e arquivamento das fichas” (BARBOSA, 1978), mas também alguns princípios utilizados até os dias de hoje. De acordo com Mey e Silveira (2009, p.71),

Cutter, considerado por Ranganathan como um gênio da biblioteconomia, não elaborou apenas um código de catalogação: sua obra é na verdade uma declaração de princípios. As regras são entremeadas com os porquês das soluções e com observações diversas, [...]. De fácil leitura, constitui um

código muito completo, incluindo a catalogação de assunto e de materiais especiais, normas de transliteração e elaboração de catálogos auxiliares.

[...] Determinou os objetivos do catálogo e ditou a visão do catalogador, que influenciam sobremaneira a Declaração dos Princípios Internacionais de Catalogação. Sua obra é um exemplo de catalogação prática adequada ao usuário. [...].

Conforme apontam Mey e Silveira (2009, p. 12), os objetivos do catálogo propostos por Cutter são aceitos até os dias de hoje, com algumas atualizações..Ele estabelecia:

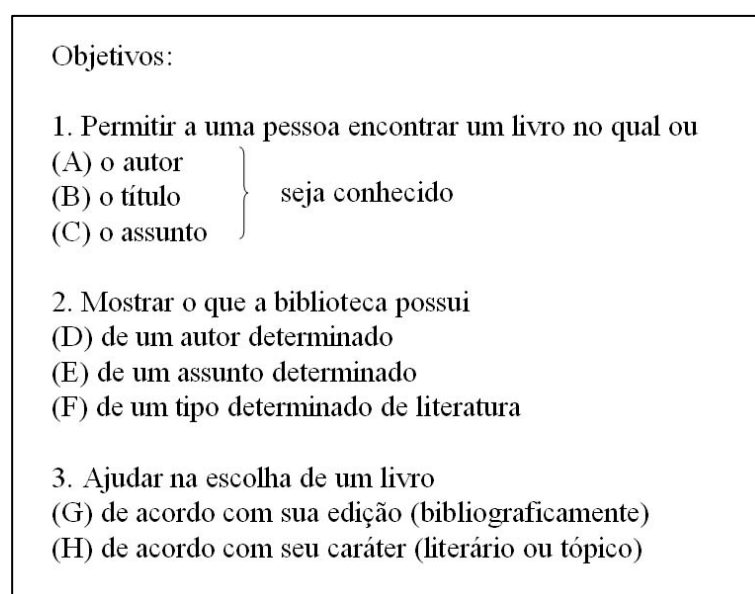


FIGURA1: Objetivos do catálogo.
Fonte: Mey, Silveira (2009, p. 12).

Pode-se considerar que a visão do catalogador definida por Cutter se refere ao estabelecimento do princípio de conveniência do público. Ele declarou que "a conveniência do público será estabelecida sempre antes da facilidade do catalogador" (CHAN, 2007, p48).

Barbosa (1978, p. 30) explica que o catálogo é o meio de comunicação entre o usuário e o acervo da biblioteca, e só poderá ser eficiente se for adotada uma padronização, entretanto, “[...] o usuário ou o público a que se destina deve ter privilégio de ser convenientemente por ele atendido, mesmo contrariando os preceitos aos quais estão ligados os catalogadores.”

Análise: os objetivos do catálogo já refletiam a preocupação de Cutter com os usuários e foram reafirmados com o estabelecimento do princípio de conveniência do usuário. O catálogo deveria refletir o que o usuário necessita e facilitar a sua busca e recuperação Essa preocupação se faz evidente nos dias de hoje com a definição dos Princípios de Paris, que refletem os objetivos e princípios idealizados por Cutter.

6.1.4 PRINCÍPIOS DE LUBETSKY

Seymour Lubetsky, bibliotecário da Library of Congress (Biblioteca do Congresso Americano), teve um importante papel em relação aos códigos de catalogação, elaborando vários estudos e publicações, tais como: *Studies of Descriptive Catalog* (1946) (estudos de catalogação descritiva); *Cataloging Rules and Principles* (1953) (regras de catalogação e princípios); *Code of Cataloging Rules* (1960) (códigos de regras de catalogação); *Principles of Cataloging* (1969) (princípios de catalogação) (CHAN, 2007).

Lubetsky foi responsável pela reavaliação dos códigos da época, “ [...] examinou a ocorrência das regras de catalogação e propôs o retorno aos princípios”; além disso, “[...] defendeu que as regras para pontos de acesso deveriam estar baseadas na “condição de autoria” e não no tipo de publicação [...]” (CHAN, 2007, p. 49).

Sobre Lubetsky, Chan (2007, p. 49, tradução nossa) destaca:

Sua insistência sobre a racionalização e aproximação dos padrões de catalogação baseados em objetivos e princípios fundamentou o desenvolvimento dos códigos de catalogação subsequentes. Dele é o mérito de transformar os códigos de catalogação “ricos em regras” para aqueles “ricos em princípios”.

A obra “*Cataloging Rules and Principles*” (Regras de Catalogação e Princípios) de Lubetsky pode ser considerada a mais importante e se divide em três partes, conforme aponta Chan (2007, p. 49):

Parte I apresenta uma análise detalhada de regras específicas da ALA (1949). Parte II começa a estudar a questão do “complexo corporativo” (novamente provendo uma análise perceptiva da confusão a respeito da autoria corporativa existente nos códigos), [...]. Parte III, “Projeto para um Código”, Lubetsky estabelece posteriormente dois objetivos: “(1) permitir ao usuário do catálogo determinar rapidamente se a biblioteca possui ou não o livro que ele quer; (2) mostrar ao usuário do catálogo, sob a forma do nome do autor, quais obras a biblioteca tem sobre um determinado autor e em quais edições ou traduções de uma determinada obra”.

Conforme apontado anteriormente, os estudos e contribuições de Lubetsky deram origem à Conferência Internacional sobre os Princípios de Catalogação, realizada com o intuito de uniformizar as regras e estabelecer princípios para a Catalogação e, a partir disso, o desenvolvimento e atualizações de vários outros códigos de catalogação (BARBOSA, 1978).

Análise: a maior contribuição de Lubetsky para a teoria da catalogação foi defender o aspecto lógico das regras e ressaltar a importância da retomada do estabelecimento de princípios que norteassem as regras, fato que estava sendo negligenciado em muitos códigos

de catalogação da época. Seu trabalho contribuiu, posteriormente, para a consolidação efetiva dos Princípios de Catalogação e a construção de códigos mais consistentes.

6.1.5 CONTRIBUIÇÕES DE GORMAN

Embora não tenha trazido contribuições para o estabelecimento de princípios, Michael Gorman é destacado nesta pesquisa na linha dos teóricos de Catalogação, por trazer significativas contribuições no desenvolvimento de um novo código de catalogação baseado nos princípios defendidos por Lubetsky e nos estabelecidos na Declaração dos Princípios Internacionais de Catalogação, mais conhecida como Princípios de Paris.

De acordo com Mey e Silveira (2009, p. 79),

[...] após estudo das informações encontradas em oito bibliografias nacionais correntes, apresentou um documento básico à RIEC, denominado International Standard Bibliographic Description [Descrição bibliográfica internacional normalizada], ou ISBD, que padronizava as informações contidas na descrição bibliográfica. Para tanto, Gorman sistematizou a ordem das informações e a pontuação utilizada antes de cada informação, de modo a tornar possível seu reconhecimento pelos computadores. [...]

A contribuição de Gorman ocorreu por meio do desenvolvimento das ISBDs e o grande avanço se constitui na sistematização das informações a serem descritas em áreas (que podemos considerar atualmente categorias de atributos), na determinação da pontuação utilizada na separação dessas áreas (ou seja, a sintaxe de codificação dos atributos) para que as informações pudessem ser distinguidas e processadas automaticamente.

Análise: com o estabelecimento das áreas, das informações a serem descritas em cada área (atributos) e com o estabelecimento da pontuação, pode-se considerar que Gorman estabeleceu a primeira estrutura descritiva do tratamento descritivo da informação ou, em outras palavras, o primeiro esquema de metadados no domínio bibliográfico, mesmo que de forma manual.

6.1.6 CONSIDERAÇÕES GERAIS SOBRE AS CONTRIBUIÇÕES PARA UMA TEORIA DA CATALOGAÇÃO

A contextualização histórica dos catálogos e da catalogação abordada no capítulo 2 possibilitou o resgate e uma breve análise dos principais aspectos e teorias desenvolvidas na catalogação, conforme apontadas neste capítulo. Assim, apresenta-se a seguir a sistematização das principais contribuições para a teoria de catalogação:

SISTEMATIZAÇÃO DAS CONTRIBUIÇÕES PARA A TEORIA DA CATALOGAÇÃO	
TEÓRICOS DA CATALOGAÇÃO	CONTRIBUIÇÕES E PRINCÍPIOS
Anthony Panizzi	<ul style="list-style-type: none"> • 91 regras de Panizzi – precursoras das regras de catalogação atuais; • A quantidade de regras refletia a especificidade de domínio; • Início do estabelecimento de padronização por meio das regras de catalogação; • Estabelecimento de alguns princípios, o objetivo do catálogo de biblioteca, a normalização da entrada principal (pessoa e entidade), a condição de obras modificadas ou adaptadas para diferenciar as originais, exigências para uniformidade na aplicação das regras de catalogação, entre outros.
Charles C. Jewett	<ul style="list-style-type: none"> • Idealizou o catálogo universal ou coletivo; • Propôs o ideal de catalogação cooperativa; • Propôs o uso de tecnologias mecânicas (estereotipia) para a facilitar a produção do catálogo.
Charles Ami Cutter	<ul style="list-style-type: none"> • Determinação das regras para catálogo dicionário; • Estabelecimento de princípios: conveniência do usuário; • Definição dos objetivos do catálogo;
Seymour Lubetsky	<ul style="list-style-type: none"> • Defendeu o aspecto lógico das regras de catalogação; • Defendeu a retomada de princípios que norteassem as regras; • Contribuições posteriores com a consolidação de princípios de catalogação e a construção de códigos mais consistentes.
Michael Gorman	<ul style="list-style-type: none"> • Sistematização da ordem das informações em áreas descritivas (determinação de categorias de atributos); • Determinação da pontuação (sintaxe de codificação dos atributos); • Estabelecimento manual do primeiro esquema de metadados do domínio bibliográfico.

QUADRO 7: Sistematização das contribuições para a teoria da catalogação.
FONTE: do autor.

Com base no exposto e na sistematização do quadro 7, pode-se verificar que de modo geral os teóricos destacados buscavam a definição de regras mais consistentes, uma unificação das regras, e principalmente o estabelecimento de princípios que norteassem essas regras e o tratamento descritivo da informação.

Um importante passo para a consolidação desses aspectos ocorreu a partir da Conferência Internacional sobre os Princípios de Catalogação e com o posterior estabelecimento da Declaração dos Princípios Internacionais de Catalogação, mais conhecido como Princípios de Paris.

Segundo Chan (2007, p. 49, tradução nossa), os Princípios de Paris determinavam os objetivos e as funções do catálogo “[...] resultantes logicamente desses objetivos e que são declarados em condições específicas e em detalhe considerável”, bem como outros princípios que norteavam a construção de regras de catalogação. O autor afirma ainda que os Princípios de Paris representaram “[...] um grande passo em direção a um acordo internacional”, pois minimizaram as principais diferenças entre os códigos de catalogação de diversos países, possibilitaram a revisão dos códigos existentes e a construção de códigos de catalogação mais consistentes (CHAN, 2007, p. 50, tradução nossa).

Assim, pode-se considerar que a Declaração dos Princípios Internacionais de Catalogação consolidou a teoria desenvolvida na catalogação subjacente à prática desenvolvida no tratamento descritivo da informação.

A partir de 2003 a Declaração dos Princípios Internacionais de Catalogação vem passando por algumas atualizações no intuito de abranger as necessidades ocasionadas pelas tecnologias no âmbito do domínio bibliográfico. A seguir, apresentam-se alguns aspectos considerados mais importantes sobre os atuais princípios de catalogação revistos.

6.2 PRINCÍPIOS E FUNDAMENTOS: Declaração dos Princípios Internacionais de Catalogação - Princípios de Paris

Assim como a catalogação vem evoluindo e aprimorando seus processos de TDI, a Declaração dos Princípios Internacionais de Catalogação necessitou acompanhar essas atualizações para refletir as diretrizes necessárias às atuais necessidades tecnológicas na representação informacional.

Procede-se neste item a uma análise dos principais aspectos dos princípios de Catalogação em relação às atuais tendências de tratamento descritivo da informação. Considera-se nesta pesquisa que as atuais tendências no TDI são representadas pelo surgimento de diferentes padrões de metadados decorrentes do crescente número de recursos informacionais digitais e diversidade de ambientes informacionais digitais.

A Declaração dos Princípios Internacionais de Catalogação ou Princípios de Paris é considerada nesta pesquisa um instrumento que determina os princípios e os fundamentos do domínio bibliográfico, além de se constituir como uma diretriz para o domínio. Uma diretriz⁷ pode ser entendida como normas de procedimento formadas por um conjunto de preceitos ou regras que determinam o modo correto a se proceder em um determinado caso (FERREIRA, 2004).

As diretrizes contempladas na Declaração Internacional de Princípios de Catalogação abarcam princípios pertencentes ao domínio bibliográfico; os objetivos e funções do catálogo, que poderiam ser entendidos como objetivos e funções do sistema de informação específico do domínio, apontam “[...] regras orientadoras que devem ser incorporadas aos códigos de catalogação em âmbito internacional [...]” e determinam orientações para viabilizar a busca e a recuperação dos recursos nos catálogos (IFLA, 2009, p. 01).

A Declaração Internacional de Princípios de Catalogação está em consonância com as atuais tendências tecnológicas para o tratamento informacional, e foi atualizada com base nos Requisitos Funcionais para Registros Bibliográficos (FRBR), instrumento que será abordado mais adiante.

Atualmente a Declaração de Princípios contempla diretrizes em relação a: abrangência dos princípios; princípios gerais do domínio; objetos do TDI (entidades, atributos e relações); objetivos e funções do catálogo (considerado um sistema de informação específico do domínio bibliográfico); descrição bibliográfica (aponta diretrizes e instrumentos); estabelece diretrizes para determinação dos pontos de acesso na recuperação; bem como fundamentos para aperfeiçoamento da busca e recuperação. Essas diretrizes estão especificadas nos seguintes tópicos:

1. Abrangência
2. Princípios Gerais
3. Entidades, Atributos e Relações
4. Objetivos e Funções do Catálogo
5. Descrição Bibliográfica
6. Pontos de Acesso
7. Fundamentos para Aperfeiçoamento da Busca (IFLA, 2009, p. 01).

Dentre esses tópicos, destacam-se inicialmente nesta pesquisa os princípios gerais e objetivos e funções dos catálogos.

⁷ “Conjunto de instruções ou indicações para se tratar e levar a termo um plano, uma ação, um negócio, etc.; diretiva”, “Norma de procedimento; diretiva” (FERREIRA, 2004). Nesta pesquisa o conceito diretriz foi definido com base nos verbetes: diretriz, norma, regra, preceito.

Os princípios gerais constituem-se como princípios essenciais do domínio bibliográfico. Foram revistos e restabelecidos de acordo com as novas necessidades e características do domínio; para isso partiu-se do modelo conceitual FRBR, que determina os requisitos funcionais do domínio.

Os princípios que norteiam o domínio bibliográfico são:

Conveniência do usuário do catálogo. As decisões relativas a descrições e formas controladas de nomes para acesso devem ser tomadas tendo o usuário em mente.

Uso comum. O vocabulário usado nas descrições e nos pontos de acesso deve ser adequado à maioria dos usuários.

Representação. Descrições e formas controladas de nomes para acesso devem ser baseadas na forma pela qual uma entidade descreve a si mesma.

Precisão. A entidade descrita deve ser retratada fielmente.

Suficiência e necessidade. Nas descrições e formas controladas de nomes para acesso, devem ser utilizados apenas os elementos necessários para o usuário e essenciais para identificar individualmente uma entidade.

Significância. Os elementos devem ser bibliograficamente significativos.

Economia. Quando existirem meios alternativos para se alcançar um objetivo, deve ser dada preferência ao meio que promova economia geral, de forma mais adequada (i.e., menor custo ou abordagem mais simples).

Consistência e padronização. Os processos de descrição e construção de pontos de acesso devem ser padronizados tanto quanto possível. Isso possibilita uma maior consistência, aumentando também a facilidade de compartilhamento de dados bibliográficos e de autoridade.

Integração. As descrições para todos os tipos de materiais e formas controladas de nomes de entidades devem ser baseadas, tanto quanto possível, em um conjunto comum de regras (IFLA, 2009, p. 02).

Já os objetivos e funções do catálogo são considerados aqui como características que norteiam um sistema específico. Os avanços da tecnologia possibilitaram a construção de modo automatizado dos catálogos e a sua disponibilização em meio digital, fato que ampliou sua função e objetivos diante das novas necessidades dos usuários. Dessa forma, o objetivo e a função do catálogo foram redefinidos, com base nos princípios gerais anteriormente citados e nos requisitos do domínio bibliográfico apontados pelo modelo conceitual FRBR.

Assim, o catálogo passa a ser considerado um instrumento importante para que o usuário possa localizar, identificar, selecionar, adquirir (ou obter) e navegar entre os registros bibliográficos. Esses objetivos e funções foram estabelecidos de acordo com as tarefas dos usuários definidas no modelo FRBR, ou seja, as tarefas que os usuários realizarão no sistema. Para melhor entendimento, destaca-se a seguir cada uma das novas funções do catálogo:

O catálogo deve ser um instrumento efetivo e eficiente que permita ao usuário:

1 Encontrar recursos bibliográficos em uma coleção como o resultado de uma busca, usando atributos e relações entre recursos:

1.1 localizar determinado recurso

1.2 localizar conjuntos de recursos representando todos os recursos que pertencem a uma mesma obra todos os recursos que incorporam uma mesma expressão todos os recursos que exemplificam uma mesma manifestação todas os recursos associados a determinada pessoa, família, ou entidade coletiva todos os recursos sobre determinado assunto todos os recursos definidos por outros critérios (língua, país de publicação, data de publicação, tipo de conteúdo, tipo de suporte etc.) geralmente como filtro secundário de busca;

2 identificar um recurso bibliográfico ou agente (isto é, confirmar que a entidade descrita corresponde à entidade desejada ou distinguir entre duas ou mais entidades com características similares);

3 selecionar um recurso bibliográfico apropriado ao usuário (isto é, escolher um recurso que esteja de acordo com as necessidades do usuário no que diz respeito ao meio, conteúdo, suporte etc. ou rejeitar um recurso não apropriado às necessidades do usuário);

4 adquirir ou **obter** acesso a um item descrito (isto é, fornecer informação que capacite o usuário a adquirir um item por meio de compra, empréstimo etc., ou acessar um item eletronicamente por meio de conexão online a uma fonte remota); ou acessar, adquirir ou obter um dado de autoridade ou bibliográfico.

5 navegar em um catálogo e além dele (isto é, por meio do arranjo lógico dos dados bibliográficos e de autoridade e de maneiras claras de se fazer esta navegação, incluindo a apresentação de relações entre obras, expressões, manifestações, itens, pessoas, famílias, entidades coletivas, conceitos, objetos, eventos e lugares). (IFLA, 2009, p. 03-04).

De acordo com a IFLA (2009, p. 02), o intuito da Declaração de Princípios é de que ela: “[...] facilite o compartilhamento internacional de dados bibliográficos e de autoridade e guie os criadores de regras de catalogação em seus esforços para desenvolver um código internacional de catalogação”. Assim, estabelece também diretrizes para outras questões importante, como, por exemplo, o objeto do processo de tratamento descritivo da informação (entidades, atributos e relacionamentos), diretrizes para a descrição bibliográfica, pontos de acesso e fundamentos para o aperfeiçoamento da busca e recuperação.

Quanto ao objeto do processo de TDI, a Declaração de Princípios estabelece uma significativa mudança, apontando que “um código de catalogação deve levar em consideração as entidades, atributos e relações tal como definido em modelos conceituais do universo bibliográfico” (IFLA, 2009, p. 03).

Esse ponto de vista muda o olhar do objeto da catalogação: o que antes era item físico, agora, sob o ponto de vista conceitual do sistema, são as entidades (obra intelectual) e seus metadados.

Considerando as entidades, atributos e relações determinadas pelo modelo conceitual do domínio, é possível criar regras mais consistentes para representar os recursos informacionais nas estruturas de descrição (esquemas) de metadados dos padrões.

Desse modo, as diretrizes para a descrição bibliográfica também se apresentam em consonância com o modelo conceitual do domínio. Em relação a esta questão, a Declaração de Princípios estabelece: a necessidade de criação de uma descrição bibliográfica para cada manifestação, incluindo atributos próprios da obra e expressão; que os dados descritivos (metadados) sejam baseados em normas internacionalmente aceitas; e que as descrições possam variar entre níveis de detalhamento, desde que seja comunicado ao usuário (respeitando o princípio de conveniência do usuário) (IFLA, 2009).

Ainda em relação à descrição bibliográfica, são apontadas também algumas diretrizes para a determinação dos pontos de acesso e suas variações, e em relação aos sistemas aponta diretrizes para a busca e aspectos de recuperação (IFLA, 2009).

Conforme pode ser observado, o intuito deste capítulo é traçar algumas considerações sobre os principais aspectos da Declaração Internacional de Princípios de Catalogação. Sendo assim, apresenta-se a seguir uma sistematização dos principais aspectos tratados na Declaração Internacional de Princípios de Catalogação, com o objetivo de categorizar os principais princípios e fundamentos norteadores do domínio bibliográfico:

SISTEMATIZAÇÃO DOS PRINCÍPIOS E FUNDAMENTOS NORTEADORES DO DOMÍNIO BIBLIOGRÁFICO	
OBJETOS DO TDI	<ul style="list-style-type: none"> • Entidades, atributos e relações (conforme estabelecidos nos modelos conceituais do domínio bibliográfico)
PRINCÍPIOS DO DOMÍNIO BIBLIOGRÁFICO	<ul style="list-style-type: none"> • Conveniência do usuário do catálogo • Uso comum • Representação • Precisão • Suficiência e necessidade • Significância • Economia • Consistência e padronização • Integração
OBJETIVOS E FUNÇÕES DO CATÁLOGO	<ul style="list-style-type: none"> • Localizar • Identificar • Selecionar • Adquirir ou obter • Navegar
DESCRIÇÃO BIBLIOGRÁFICA	<ul style="list-style-type: none"> • Elaborada para cada manifestação de uma entidade (recurso informacional) • Pode haver descrições em diversos níveis de detalhamento • A descrição deve estar baseada em norma internacionalmente aceita (ISBD - International Standard Bibliographic Description)
PONTOS DE ACESSO	<ul style="list-style-type: none"> • Estabelece os princípios gerais para determinação dos pontos de acesso e determina outros princípios para a escolha dos pontos de acesso, para os pontos de acesso autorizados, para nomes variantes e formas variantes de nomes
FUNDAMENTOS E APERFEIÇOAMENTO PARA BUSCA E RECUPERAÇÃO	<ul style="list-style-type: none"> • De modo geral determina os pontos de acesso essenciais e pontos de acesso adicionais para busca; • Determina o uso de uma ordem lógica para visualização dos registros com o mesmo ponto de acesso aos usuários do catálogo

QUADRO 8: Sistematização dos princípios e fundamentos norteadores do domínio bibliográfico.

FONTE: do autor.

Entende-se que uma disciplina, como qualquer ramo do conhecimento científico, é formada com base em fundamentos ou em um conjunto de princípios básicos e preceitos de um domínio. Dessa forma, a teoria de uma disciplina é composta pelo conjunto de princípios ou fundamentos, que foram construídos com base em um conjunto de conhecimentos advindos a partir de uma prática (FERREIRA, 2004).

Considera-se nesta pesquisa a existência de uma teoria para a Catalogação, desenvolvida por teóricos e instituições interessadas em promover uma uniformidade e uma padronização no TDI. Essa teoria, desenvolvida a partir da análise do processo prático de

TDI, foi estabelecida por meio dos princípios que fundamentam a disciplina e que passam a ser refletidos nos instrumentos de representação no domínio bibliográfico.

Acredita-se que os princípios gerais apontados anteriormente, bem como os outros aspectos considerados, constituem-se como fundamentos para o domínio bibliográfico, auxiliado na construção consistente de catálogos e descrições bibliográficas (consolidadas por meio de códigos e regras). Em outras palavras, considera-se que a Declaração Internacional de Princípios de Catalogação constitui-se como um fundamento para o domínio bibliográfico.

6.2.1 INSTRUMENTOS DE REPRESENTAÇÃO: modelos conceituais, esquemas de codificação, padrões de metadados

Conforme apontado no capítulo anterior, a construção de representações é algo complexo, envolve aspectos representacionais, principalmente em relação à determinação padronizada de metadados, seus valores e sua contextualização nos sistemas; e aspectos tecnológicos, relacionados a padrões técnicos e de infraestrutura nos sistemas de informação que consolidam a existência dos metadados nos sistemas (MÉNDEZ RODRÍGUEZ, 2002).

A união desses dois aspectos promove a determinação consistente de metadados, contudo, conforme já salientado, é importante que sejam consideradas as características, princípios e fundamentos estabelecidos em cada domínio, pois isso norteará a construção adequada de representações tornando-as específicas de um domínio, lembrando que uma representação adequada é considerada nesta pesquisa uma representação padronizada e que corresponde de modo exato com seu objeto, ou seja, uma representação concisa e coerente.

É relevante reafirmar que o processo de catalogação no domínio bibliográfico consiste na criação e padronização de metadados específicos para o tratamento descritivo informacional. Para que isso seja feito de modo consistente, a determinação e a criação dos metadados dependem da aplicação de instrumentos, por isso a importância de se estabelecer princípios em um domínio para nortear a construção apropriada desses instrumentos de representação.

Pela história da catalogação apontada no capítulo 2, pode-se observar que o domínio bibliográfico sempre buscou desenvolver instrumentos de representação. Entretanto, diante das atuais mudanças causadas pelos avanços tecnológicos pode-se considerar que os métodos

atuais de tratamento descritivo da informação (TDI) são adequados para a construção padronizada de metadados no domínio bibliográfico?

Para responder a essa questão, busca-se neste item categorizar os instrumentos de representação do domínio bibliográfico necessários para promover a determinação e a criação dos metadados e suas representações, tendo com base os aspectos tecnológicos e representacionais abordados no capítulo anterior.

6.2.1.1 MODELO CONCEITUAL

De acordo com a Declaração dos Princípios Internacionais de Catalogação, os modelos conceituais do domínio bibliográfico são: FRBR - *Functional Requirements for Bibliographic Records* (Requisitos Funcionais para Registros Bibliográficos), FRAD - *Functional Requirements for Authority Data* (Requisitos Funcionais para Dados de Autoridade), FRSAD - *Functional Requirements for Subject Authority Data* (Requisitos Funcionais para Dados de Autoridade de Assunto) (IFLA, 2009).

Muito estudado atualmente, o FRBR constitui-se como um modelo conceitual que oferece um modo de entender as relações em uma representação ou registro bibliográfico.

De acordo com a IFLA (2005), seu propósito é “recomendar um nível básico de funcionalidade e uma exigência básica dos elementos para os registros criados pelas agências bibliográficas nacionais”.

Para Moreno (2005, p. 19, grifo do autor), o modelo FRBR pode ser explicado como “[...] uma recomendação para **reestruturar** os registros bibliográficos de maneira a refletir a estrutura conceitual de **buscas** de informação [...]”. A estrutura conceitual de buscas de informação é determinada pelas tarefas dos usuários.

Segundo a IFLA (1998), o FRBR representa uma tentativa inicial de se estabelecer critérios para a compreensão e desenvolvimento de registros bibliográficos, e se constitui como um modelo que atua no nível conceitual, estabelecendo relações entre entidades, atributos e relacionamentos.

Os FRBR foram estabelecidos com base nos princípios de catalogação, nos códigos de representação e padrões ou formatos de metadados da Biblioteconomia e, sendo um modelo conceitual, auxilia a modelagem do banco de dados, refletindo o universo bibliográfico; além

disso, define por meio de suas entidades, atributos e relacionamentos os requisitos necessários para a construção de representações bibliográficas.

Portanto, os FRBR constituem-se como o mais novo modelo de requisitos para a modelagem de catálogos e bancos de dados a partir da construção normalizada de representações bibliográficas.

O modelo conceitual FRAD e o modelo conceitual FRSAD estão sendo estabelecidos para auxiliar a estruturação das entidades, atributos e relações no que diz respeito a dados de autoridade e assunto.

Os modelos conceituais são baseados na metodologia de entidade/relacionamento (E-R). Buscam identificar os requisitos funcionais do sistema, tendo como base as tarefas dos usuários, ou seja, quais tarefas o usuário realizará naquele sistema, e a partir daí mapeiam as entidades, atributos e relações em um domínio.

De modo geral, os modelos conceituais do domínio bibliográfico buscam definir uma estrutura clara da relação entre os registros bibliográficos e as necessidades dos usuários (identificadas nas tarefas dos usuários). Os modelos conceituais mapeiam os requisitos do sistema e auxiliam o aperfeiçoamento de instrumentos de descrição (esquemas, códigos, regras), na definição mais clara das entidades pertencentes a um grupo, na representação derivada dos instrumentos e da estrutura do sistema (catálogo).

Os modelos conceituais do domínio bibliográfico podem ser considerados instrumentos que definem conceitualmente um sistema de informação (o catálogo) e os componentes do sistema (entidades, atributos e relações, ou ainda recursos informacionais e seus metadados). Definem mais claramente a estrutura de relacionamentos auxiliando a reavaliação da estrutura dos instrumentos de representação (esquemas, códigos) e posteriormente na definição de regras mais consistentes de representação.

A determinação dos requisitos do sistema por meio do mapeamento das necessidades dos usuários (tarefas dos usuários) e dos objetos do domínio (entidades, atributos e relacionamento) possibilitam a construção de regras mais consistentes, pois se conhece a estrutura lógica e os relacionamentos entre atributos e entidades.

Os modelos conceituais do domínio bibliográfico são instrumentos que convergem aspectos tecnológicos e representacionais, determinam de forma lógica e conceitual as entidades, seus atributos e os relacionamentos existentes entre elas, no intuito de determinar no ambiente tecnológico uma estrutura clara dos requisitos do domínio bibliográfico a serem contemplados no sistema.

Assim, os modelos conceituais confirmam a integração estratégica entre as tecnologias de informática e o tratamento descritivo da informação e, neste caso, as tecnologias de informática atual como agentes determinantes de um instrumento conceitual para o domínio bibliográfico.

6.2.1.2 ESQUEMAS DE CODIFICAÇÃO

Conforme afirmado no capítulo anterior, os esquemas são instrumentos relacionados aos aspectos representacionais para determinação dos metadados. São subdivididos em esquemas de metadados e esquemas de codificação, ambos utilizados para proporcionar maior consistência na representação.

Os esquemas de metadados são o conjunto de elementos descritivos (metadados ou atributos codificados) componentes de um padrão de metadados. Serão abordados no item seguinte.

Os esquemas tratados aqui referem-se aos esquemas de codificação de dados, que determinam a consistência dos dados nas estruturas descritivas ou esquemas de metadados do padrão. Os esquemas de codificação correspondem a esquemas externos à estrutura de descrição do padrão e são representados no tratamento descritivo da informação pelos códigos de catalogação⁸.

Como exemplos de esquemas do domínio bibliográfico podem ser citados: o código AACR2 (*Anglo-American Cataloging Rules – Código de Catalogação Anglo-Americano*, 2ª edição), as ISBD(s) (*International Standard Bibliographic Description – Padrão Internacional de Descrição Bibliográfica*) e o novo código RDA (*Resource Description and Access – Descrição do Recurso e Acesso*).

O AACR2 constitui-se como um código que apresenta um conjunto de regras e normas para a construção e padronização da representação bibliográfica de diversos recursos informacionais. De modo geral, suas regras definem a forma de transcrição dos valores, onde podem ser encontrados, o que deve ser feito em casos especiais, quais valores devem estar

⁸ Coleção de regras e preceitos (FERREIRA, 2004). O Código de catalogação estabelece os preceitos e regras para a representação padronizada do recurso, em outras palavras, estabelece o modo de proceder para a construção de uma descrição formal e consistente de uma entidade.

contidos em um determinado elemento de descrição, entre outras regras. Deste modo, o AACR2 normaliza a representação (valores) no padrão de metadados MARC 21.

O RDA vem ganhando destaque por estar sendo construído de modo alinhado ao modelo conceitual FRBR e, de acordo com a explicação de Gómez (2007), está organizado da seguinte forma: divide-se em duas grandes partes, descrição e controle de pontos de acesso. Na primeira parte há outra divisão, uma para a descrição e a outra para as relações entre as entidades. Contudo, o diferencial é que o RDA não está organizado por tipo de material, mas pelos distintos atributos utilizados para a descrição de cada recurso (título, variações do título, indicação de responsabilidade etc).

Além disso, é possível estabelecer diferentes níveis de detalhe na catalogação, ou seja, escolher diferentes níveis de especificidade sem, contudo, perder a formalidade do esquema de descrição e as regras para a construção da representação no formato.

Não se sabe até o momento qual será seu alcance e aceitação na comunidade bibliotecária, assim não é objetivo aqui traçar todas as características do novo código, mas apenas destacá-lo como um novo esquema para a construção de representações padronizadas e que está em consonância com o modelo conceitual da área.

De acordo com Gómez (2007), a aplicação do modelo conceitual e do novo esquema de representação proporcionará um ótimo aproveitamento das tecnologias (GÓMEZ, 2007).

Porém, destaca-se que, independente de qual esquema se consolidará na área, o importante é seu uso, pois garante maior padronização da representação nos padrões de metadados da área.

Esses esquemas, em geral, apresentam regras que padronizam a representação dos seguintes aspectos (MILLER, 2008):

- Estrutura dos elementos: os esquemas de codificação definem os metadados pertencentes à estrutura descritiva do esquema de metadados do padrão. Essa definição é fornecida pelas áreas e subáreas estabelecidas nas ISBDs e no AACR e determinam a definição dos elementos do esquema do padrão MARC 21;
- Conteúdo dos dados: cada área e subárea apresenta regras específicas para representação nos elementos de metadados; além disso, o conteúdo dos dados pode ser melhor definido com as diretrizes do domínio e guias de boas práticas;
- Codificação dos dados: as áreas, as subáreas, a pontuação e as regras dos códigos estabelecem a ordem lógica e padronizada de codificação dos dados a serem legíveis por máquina; porém, o intercâmbio dos dados é proporcionado por outros esquemas de

intercâmbio de dados, como o esquema da linguagem XML ou ANSI/NISO Z39.50 e a ISO 2709;

- Apresentação dos dados: refere-se à padronização na apresentação (visualização) da representação aos usuários; os esquemas de codificação ISBD e AACR determinam a apresentação pela ordem das áreas e a pontuação. Na maioria dos casos a apresentação dos dados é dada pelo significado dos atributos e não pelo atributo codificado;
- Valores dos dados: referem-se aos esquemas de codificação que determinam os valores possíveis a serem admitidos nos espaços do valor dos elementos de metadados (ex.: tesouros, vocabulários controlados etc).

É importante destacar que o aprimoramento dos esquemas de representação da área reflete a necessidade de consonância com o modelo conceitual, pois os esquemas é que proporcionarão a normalização necessária aos padrões ou formatos de metadados da área.

6.2.1.3 PADRÃO DE METADADOS

Conforme se evidenciou em capítulos anteriores, os padrões de metadados são compostos pelo conjunto de elementos descritivos, metadados codificados que descrevem um recurso informacional. Esses metadados são metodologicamente construídos com base nos princípios estabelecidos no domínio, nos esquemas de codificação (códigos que estabelecem regras) e, portanto, são estabelecidos previamente. Assim, os padrões de metadados consolidam em uma estrutura automatizada os metadados presentes nos esquemas de codificação manual do domínio bibliográfico.

Pode-se dizer que o uso apropriado de um padrão de metadados proporciona uma descrição normalizada e a construção de representações concisas, claras e de qualidade, facilitando o intercâmbio de informações, a interoperabilidade entre sistemas e a recuperação da informação. Para isso, é necessário que seja considerado o tipo de ambiente informacional e, principalmente, qual o princípio que norteia o padrão a ser utilizado.

A autora Méndez Rodríguez (2002) ressalta um ponto significativo: afirma que é importante que os metadados não sejam iniciativas caprichosas de áreas científicas distintas, devem responder às necessidades de recuperação da informação e constituir-se como padrões reconhecidos e compartilhados nessas áreas para facilitar o intercâmbio de informações. Isso

justifica os esforços, em diferentes domínios, para a construção de instrumentos norteadores e de padronização dos metadados.

Com base no exposto, considera-se que a construção de representações adequadas está pautada no uso de padrões de metadados correspondentes ao tipo de ambiente informacional e de ferramentas de normalização dos metadados, tais como os modelos conceituais e os esquemas de codificação. Assim, apresenta-se a seguir a categorização dos instrumentos de representação utilizados no TDI no domínio bibliográfico:

INSTRUMENTOS DE REPRESENTAÇÃO NO TDI	
MODELO CONCEITUAL	
REQUISITOS FUNCIONAIS PARA REGISTROS BIBLIOGRÁFICOS	<ul style="list-style-type: none"> FRBR - <i>Functional Requirements for Bibliographic Records</i>
REQUISITOS FUNCIONAIS PARA DADOS DE AUTORIDADE	<ul style="list-style-type: none"> FRAD - <i>Functional Requirements for Authority Data</i>
REQUISITOS FUNCIONAIS PARA DADOS DE AUTORIDADE DE ASSUNTO	<ul style="list-style-type: none"> FRSAD - <i>Functional Requirements for Subject Authority Data</i>
ESQUEMAS DE CODIFICAÇÃO	
ESQUEMAS DE CODIFICAÇÃO DE CONTEÚDO	<ul style="list-style-type: none"> Padrões de estrutura de dados: Esquema de metadados: estrutura descritiva em diretórios, campos, subcampos, código de subcampo, indicadores
	<ul style="list-style-type: none"> Padrões de intercâmbio de dados: Protocolo ANSI/NISO Z39.50; Norma ISO 2709
ESQUEMAS DE CODIFICAÇÃO DE VALORES	<ul style="list-style-type: none"> Padrões de conteúdo de dados: Padrões ou códigos elaborados para o TDI. Ex.: AACR, ISBDs
	<ul style="list-style-type: none"> Padrões de valores de dados: Tesouros, listas de autoridade, léxicos, esquemas de classificação, vocabulários controlados
PADRÃO DE METADADOS	
<ul style="list-style-type: none"> MARC 21 e desdobramentos, incluindo a versão em MARC XML 	

QUADRO 09: Categorização dos instrumentos de representação no TDI no domínio bibliográfico.

FONTE: do autor.

É importante destacar que maiores detalhes sobre os padrões de metadados, modelos conceituais e esquemas de codificação já foram abordados respectivamente nos seguintes itens: 4.2, 5.1.1 e 5.2. Destaca-se ainda que o quadro 06 apresenta uma sistematização mais detalhada dos esquemas de codificação, tanto para o domínio Web como para o domínio bibliográfico. Assim, considera-se o quadro 6 nesta pesquisa como uma especificação do instrumento esquema de codificação apontado na categorização do quadro 09.

Conforme já abordado em 4 e 4.2, o padrão de metadados do domínio bibliográfico apresenta uma estrutura rica em detalhes descritivos, proporcionando maior especificidade na representação. É possível que instituições distintas utilizem e determinem diferentes níveis de especificidade para representar um recurso informacional; contudo, os princípios que norteiam o estabelecimento de uma estrutura coerente de descrição para o domínio bibliográfico já foram estabelecidos (conforme apontado no item 6.2), conferindo ao padrão maior uniformidade.

Além disso, acredita-se também que o uso conjunto dos instrumentos apontados na categorização apresentada no quadro 09 garante a construção de representações adequadas e de qualidade, facilitando a busca e a recuperação baseada em metadados.

Após a categorização dos instrumentos utilizados para a construção de representações no domínio bibliográfico, apresenta-se a seguir um quadro com a sistematização geral dos componentes do TDI:

SISTEMATIZAÇÃO GERAL DOS COMPONENTES DO TDI	
CONTRIBUIÇÕES PARA UMA TEORIA DA CATALOGAÇÃO	Princípios de Panizzi Princípios de Jewett Princípios de Cutter Princípios de Lubetsky Contribuições de Gorman
PRINCÍPIOS E FUNDAMENTOS	Declaração dos Princípios Internacionais de Catalogação - Princípios de Paris
INSTRUMENTOS DE REPRESENTAÇÃO	Modelo Conceitual Esquemas de codificação Padrão de metadados

QUADRO 10: SÍNTESE GERAL DOS COMPONENTES DO TDI.

FONTE: do autor.

O intuito da sistematização apontada no quadro 10 é demonstrar a consonância entre a teoria desenvolvida na catalogação para o processo de TDI, que é refletida por meio dos princípios e fundamentos da disciplina, e os instrumentos necessários para a construção de representações adequadas e de qualidade nos padrões de metadados do domínio bibliográfico.

6.3 TENDÊNCIAS ATUAIS NO MÉTODO DE TRATAMENTO DESCRITIVO DA INFORMAÇÃO NO DOMÍNIO BIBLIOGRÁFICO

Diante das mudanças decorrentes dos avanços tecnológicos ocorridos a partir da década de 1990 do século XX como, por exemplo, novos ambientes informacionais em meio digital, quantidade crescente de novos tipos de recursos informacionais e o surgimento dos metadados, o domínio bibliográfico vem enfrentando uma nova realidade: a necessidade de uma reavaliação de suas teorias, princípios e fundamentos e instrumentos de construção de representações.

É importante destacar que essas são mudanças complexas e que não ocorrem de um momento para outro. Além disso, constituem-se em mudanças inter-relacionadas, pois devem aliar a teoria e a prática do tratamento descritivo da informação estabelecidos no domínio bibliográfico (BIANCHINI; GUERRINI, 2009).

As tendências atuais no método⁹ de tratamento descritivo da informação no domínio bibliográfico vêm sendo alteradas pela reavaliação dos princípios, fundamentos e instrumentos de representação informacional.

São duas as etapas nos métodos para o TDI. A primeira consiste na determinação prévia dos metadados por meio da abordagem manual de geração de metadados, baseada na construção consistente de princípios, fundamentos e instrumentos de descrição e representação (ZENG, QIN, 2008). A segunda consiste na aplicação desses princípios, fundamentos e instrumentos em estruturas padronizadas de descrição (padrões de metadados).

O momento atual ainda se encontra na primeira etapa, com poucas aplicações efetivas na mudança prática dos métodos de TDI.

Os autores Bianchini e Guerrini (2009) observam que a complexidade dessas mudanças consiste na reestruturação das bases teóricas e práticas pautadas nos FRBR – Requisitos Funcionais para Registros Bibliográficos. Destacam algumas outras preocupações, ressaltadas aqui para destacar sua importância e estabelecer uma reflexão:

Por muitas razões, este é um processo muito difícil. Não somente cada mudança é complexa, mas muitas mudanças prosseguem simultaneamente e nenhuma delas pode contar com qualquer uma das outras, porque tudo está mudando ao mesmo tempo. Complexidade é maior do que seria para cada processo único. Por essas razões, cada mudança tem que ser considerada

⁹ Entende-se o método de TDI como o modo de proceder para que ocorra o processo de representação.

com um nível crescente de cuidado (BIANCHINI; GUERRINI, 2009, p. 106).

Os avanços tecnológicos surgidos exigiram uma mudança das práticas de representação no domínio bibliográfico e como consequência exigiram também uma reavaliação e restabelecimento das bases teóricas. Era preciso reformular as bases teóricas para refletir as tecnologias de informática subjacentes à teoria e à prática no tratamento descritivo da informação, construindo assim teorias, princípios, fundamentos, e instrumentos em consonância com as tecnologias, e de modo padronizado. Por esse motivo, considera-se nesta pesquisa a importância de se conhecer as bases teóricas da Catalogação, pois se constituem como princípios para o tratamento descritivo da informação no domínio bibliográfico no contexto dos avanços tecnológicos.

De acordo com Bianchini e Guerrini (2009, p. 106),

O universo bibliográfico pode ser gerenciado por uma contínua relação entre teoria e prática. Nós temos agora uma forte influência de um modelo teórico (FRBR) em todas as nossas práticas, mas não há nenhuma interação produtiva entre aquela teoria e a prática de catalogação. Nós acreditamos que deve haver uma pausa fundamental com a prática passada para abrir completamente um espaço para novos modelos e instrumentos. Porém, a troca do passado para o futuro não deve nos impedir de pensar no presente também. Particularmente, nós precisamos urgentemente chegar a um acordo e uma definição das corretas relações entre FRBR, ISBD, e códigos nacionais, multinacionais e internacionais, chefiados pelo RDA (BIANCHINI; GUERRINI, 2009, p. 106).

Com base nessa afirmação, apontam-se aqui algumas análises tendo como base os objetivos propostos para a solução dos problemas de pesquisa. As considerações dessas análises tornam-se importantes para entender o contexto atual e as consequências dessas mudanças.

A chave para essas mudanças recai nas tecnologias de informática instauradas nos processos de tratamento descritivo da informação.

Só foi possível realizar muitos dos princípios idealizados no início da consolidação da Catalogação com o surgimento da informática e sua utilização nos processos de representação. Um exemplo são os ideais de catalogação cooperativa iniciados por Jewett. É importante destacar que as tecnologias exigiam uma padronização, mas o tratamento descritivo da informação só iria se efetivar em grande escala com a implantação das tecnologias, ainda mais com o surgimento de ambientes informacionais digitais.

Presencia-se atualmente a convergência desses dois aspectos: o representacional, marcado pelo desenvolvimento e aprimoramento de padrões para codificação dos dados e

consequentemente o aprimoramento de teorias, princípios, fundamentos e instrumentos para o tratamento descritivo da informação; e o aspecto tecnológico, marcado pelo desenvolvimento e aprimoramento de ferramentas tecnológicas para estruturação dos recursos informacionais, métodos mais eficientes para a construção de bancos de dados em diferentes ambientes informacionais digitais.

A diferença no TDI entre o momento atual, marcado pelo surgimento dos metadados, e os períodos anteriores a 1990, é que as tecnologias de informática deixam de ser apenas instrumentos facilitadores do processo de catalogação, e passam a ser agentes determinantes de mudança. Essas mudanças causadas pelas tecnologias incidiram sobre a necessidade de uma revisão e aprimoramento das teorias, dos princípios, dos fundamentos e dos instrumentos de representação no domínio bibliográfico, no intuito de serem melhor estabelecidos em meio digital. Em outras palavras, as tecnologias concretizam a existência do TDI no ambiente digital, consolidando a existência dos metadados do domínio bibliográfico em uma estrutura passível de ser processada automaticamente.

É importante ressaltar também que, para o desenvolvimento de soluções tecnológicas consistentes, as teorias, os princípios, os fundamentos e os instrumentos de tratamento descritivo da informação constituem-se como elementos orientadores no desenvolvimento dessas soluções tecnológicas em relação aos processos de representação, armazenamento e recuperação da informação no domínio bibliográfico.

Assim, retoma-se a questão da integração entre os aspectos tecnológicos e representacionais e destaca-se ainda a importância dos tradicionais princípios, fundamentos e instrumentos de representação da Catalogação como metodologias de padronização do TDI nos formatos de metadados da área de Ciência da Informação.

Destaca-se também que esses mesmos princípios, fundamentos e instrumentos de representação se constituem como ferramentas atuais para o tratamento descritivo da informação diante das novas exigências trazidas pela implementação das tecnologias de informática, sem a necessidade de uso de outros padrões que não sejam os já disponíveis na Ciência da Informação.

7 CONSIDERAÇÕES FINAIS

Conforme observado no decorrer da pesquisa, as tecnologias de informática atuam significativamente sobre as mudanças ocorridas no tratamento descritivo da informação (TDI). As tecnologias consolidaram a existência dos aspectos representacionais em meio digital, enquanto os aspectos representacionais padronizavam e proporcionavam maior consistência às soluções tecnológicas propostas.

Dois aspectos se destacam como consequência das mudanças tecnológicas nos processos de TDI: os aspectos relacionados à concepção dos sistemas, exigindo uma padronização dos dados legíveis por máquina, e os aspectos relacionados à inserção das tecnologias de informática no processo de tratamento descritivo da informação e criação de representações, exigindo uma padronização dos dados legíveis por humanos.

Neste contexto, a construção de metadados e a aplicação de padrões de metadados passam a ser um fator fundamental nos sistemas. É importante ressaltar que os constantes avanços tecnológicos vêm exigindo uma reavaliação e um aprimoramento dos processos de tratamento descritivo da informação, principalmente em domínios específicos como o bibliográfico. Assim, é importante conhecer o contexto no qual se insere o tratamento descritivo da informação, para que se possa entender as necessidades atuais de representação da informação e as mudanças no processo de tratamento descritivo da informação.

O uso intensivo de tecnologias vem ocasionando novas mudanças no tratamento descritivo da informação que podem ser constatados pela variedade de padrões de metadados surgidos nos últimos tempos.

Assim, a principal questão que motivou esta pesquisa foi: como construir representações adequadas e padronizadas que garantam a unicidade e a recuperação eficiente dos recursos informacionais na área de Ciência da Informação?

Conforme se buscou demonstrar, entende-se que a área da Ciência da Informação possui um caráter interdisciplinar e que por isso utiliza-se, em parte, dos processos de tratamento descritivo da informação (TDI) desenvolvidos na disciplina de Catalogação na área de Biblioteconomia para solucionar os problemas inerentes a essa ciência: tornar acessível e disponível a variedade crescente de recursos informacionais, bem como promover uma melhora na recuperação desses recursos.

Além disso, afirma-se que a Ciência da Informação dispõe de uma teoria e uma metodologia para o tratamento descritivo da informação (TDI) advinda da integração estratégica entre o processo de TDI desenvolvido no domínio bibliográfico e as tecnologias, em especial as tecnologias de informática, desenvolvidas ao longo do tempo para a realização mais eficiente dos processos de representação e recuperação.

Entretanto, percebe-se que os métodos de tratamento descritivo da informação no domínio bibliográfico vêm sendo questionados em relação a sua validade diante das necessidades tecnológicas atuais, levando em alguns casos a uma visão equivocada sobre o uso de padrões de metadados emergentes como sendo ideais para solucionar as questões de representação em ambientes digitais.

Nesta pesquisa os metadados são definidos como atributos referenciais codificados que representam um recurso informacional de forma única. Dessa forma, qual sua relação com o TDI? Conforme esta pesquisa, os metadados vêm sendo produzidos e padronizados há muito tempo na área de biblioteconomia, em especial nos processos de TDI. Contudo, por se tratar de um domínio específico, os metadados produzidos seguem princípios particulares pertencentes ao processo de TDI do domínio bibliográfico.

Observa-se que a diversidade de padrões de metadados criados não atende satisfatoriamente às necessidades informacionais já estabelecidas em um domínio específico como o bibliográfico, pois apresenta princípios diferentes para a criação de seus padrões. Assim, ressalta-se a importância de se utilizarem padrões de metadados que correspondam à necessidade do domínio, ou seja, é necessário que se conheça o princípio de criação do padrão de metadados para que o padrão possa atender adequadamente às necessidades informacionais do domínio.

Diante disso, quais as teorias, os princípios, os fundamentos e os instrumentos que norteiam a construção de representações no domínio bibliográfico? Qual padronização seguir para proporcionar uma determinação adequada de metadados e conseqüentemente a construção consistente de representações? Essas são algumas das questões suscitadas pelo problema de pesquisa, às quais se objetivou responder.

Com a verificação das teorias, dos princípios, dos fundamentos, dos métodos e dos instrumentos do tratamento descritivo da informação no domínio bibliográfico, desenvolvida no decorrer da pesquisa e em especial no capítulo 6, foi possível concluir que:

- Os metadados são elementos inerentes aos sistemas de informação e, portanto, elementos inerentes aos catálogos;

- Ao longo do tempo, a catalogação vem desenvolvendo e aprimorando métodos do TDI, bem como instrumentos de representação, princípios e fundamentos que consolidam a prática do processo de representação da informação;
- Os métodos de TDI foram desenvolvidos buscando a integração estratégica com as tecnologias, em especial as tecnologias de informática;
- Inicialmente as tecnologias eram empregadas como facilitadoras do processo de TDI e constituíam-se como iniciativas mecânicas. Com o desenvolvimento das tecnologias de informática, o processo de TDI passou a ser automatizado, e as tecnologias passaram a ser agentes causadores de mudanças, exigindo uma adaptação dos instrumentos de tratamento informacional desenvolvidos no domínio bibliográfico para o contexto tecnológico;
- Com o surgimento dos padrões de metadados, passa a existir uma variedade de tipos e níveis de padrões, bem como de representações. Quanto mais específico e especializado for o ambiente informacional do domínio, maior será a necessidade de um padrão de metadados altamente estruturado, para garantir a especificidade. Nesse sentido, é necessário conhecer o propósito do padrão, pois ele estará diretamente relacionado com o princípio de sua criação e com os princípios que o fundamentam no domínio para a criação de metadados e representações consistentes;
- Atualmente duas tendências são refletidas nos sistemas de informação em diferentes domínios: tendência tecnológica, caracterizada pelo desenvolvimento e aplicação de ferramentas tecnológicas que melhoram os processos de tratamento, busca e recuperação em sistemas de informação; e a tendência representacional, caracterizada pelo desenvolvimento e aplicação de padrões de metadados e esquemas de codificação para a determinação padronizada dos metadados;
- Como a eficiência do sistema depende da determinação e construção consistente de metadados, o domínio bibliográfico passa por uma reavaliação de suas teorias, princípios e instrumentos de representação;
- Com os avanços tecnológicos, é cada vez mais necessário que o catálogo execute tarefas lógicas de busca, recuperação e associação entre metadados. Com a mudança de contexto, mudam os princípios que norteiam o domínio, exigindo uma reavaliação das teorias, dos fundamentos e dos instrumentos;
- Torna-se cada vez mais importante a construção consistente de metadados; assim, devem-se observar aspectos tecnológicos e representacionais em sua determinação. Os aspectos tecnológicos destacados nesta pesquisa referem-se à determinação dos metadados e sua

estruturação nos sistemas de informação, em especial à determinação da modelagem conceitual, dos requisitos funcionais do sistema e dos esquemas de metadados do padrão (determinação da padronização dos dados legíveis por máquina). Os aspectos representacionais referem-se ao uso de esquemas de codificação externos ao padrão de metadados e que, de modo geral, determinam a ordem lógica de descrição e os valores possíveis de serem admitidos nos metadados (determinação da padronização dos dados legíveis por humanos);

- No domínio bibliográfico, as tecnologias de informática interagem estrategicamente com as teorias, com os princípios, com os fundamentos do TDI, atuando como causadora de mudanças nos métodos de TDI;
- Assim, as tecnologias de informática vêm contribuindo para uma reavaliação e aprimoramento da teoria, dos princípios, dos fundamentos, dos métodos e instrumentos de representação do domínio bibliográfico, no intuito de tornar o TDI mais consistente diante das novas necessidades dos sistemas de informação;
- Com o desenvolvimento da pesquisa foi possível constatar que as teorias, os princípios, os fundamentos e os instrumentos de TDI no domínio bibliográfico constituem-se como elementos orientadores no desenvolvimento de soluções tecnológicas nos processos de representação, armazenamento e recuperação da informação no domínio bibliográfico;
- Dessa forma, considera-se nesta pesquisa que os aspectos representacionais não se constituem nos esquemas de codificação apenas como ferramentas para a padronização, mas como instrumentos e diretrizes fundamentais para a determinação dos metadados e seus valores (representação) em um esquema de metadados do padrão, bem como as formas aceitas para intercâmbio de dados de modo consistente.
- Considera-se que os princípios, os fundamentos e os instrumentos do domínio bibliográfico constituem-se como ferramentas atuais para o TDI no domínio, pois integram-se estrategicamente com as tecnologias de informática na busca de soluções para o tratamento descritivo da informação, fato que comprova a hipótese de pesquisa.

Assim, comprova-se a tese de que a integração estratégica entre as tecnologias de informática e os métodos utilizados nos processos de tratamento descritivo da informação (TDI) na Ciência da Informação consolidam a construção padronizada e consistente de metadados, proporcionando uma representação que garanta a unicidade e a recuperação dos recursos em diferentes ambientes informacionais digitais.

Considera-se nesta pesquisa que os metadados desenvolvidos e padronizados no domínio bibliográfico constituem-se como elementos essenciais do processo de TDI, não havendo necessidade de uso de outros padrões que não sejam os já desenvolvidos no domínio.

Como conclusão, ressalta-se a importância da interação entre teoria e prática nos processos de TDI, em especial no desenvolvimento de aplicações práticas dos instrumentos de representação desenvolvidos atualmente, para que possa ser aprimorado o tratamento descritivo da informação, diante dos avanços tecnológicos atuais.

REFERÊNCIAS

ALVES, R. C. V. et.al. Ciência da Informação, Ciência da Computação e Recuperação da Informação: algumas considerações sobre os métodos e tecnologias da informação utilizados ao longo do tempo. *Revista Eletrônica Informação e Cognição*, v.6, n.1, p.28-40, 2007.

Disponível em:

<<http://www.portalppgci.marilia.unesp.br/reic/include/getdoc.php?id=210&article=60&mode=pdf>>. Acesso em: 29 set. 2007.

ALVES, R. C. V. *Web Semântica: uma análise focada no uso de metadados*. 2005, 180f. Dissertação (Mestrado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.

BACA, M. (Ed.). *Introducción a los metadatos vías a la información digital*. Traducido al español por Marisol Jacas-Santoll. Los Angeles, CA: J. Paul Getty Trust, 1998. 43 p.

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

BARRETO, C. M. *Modelo de metadados para a descrição de documentos eletrônicos na web*. 1999. 189 f. Dissertação (Mestrado em Ciências em Sistemas de Computação)-Instituto Militar de Engenharia, Rio de Janeiro, 1999. Disponível em:

<<http://ipanema.ime.eb.br/~de9/teses/1999/>>. Arquivo: cássia.zip. Acesso em: 13 set. 2001.

BIANCHINI, C.; GUERRINI, M. From bibliographic models to cataloging rules: remarks on FRBR, ICP, ISBD, and RDA and the relationships between them. *Cataloging & Classification Quarterly*, [s. l.], v. 47, n. ?, p. 105-124, 2009.

BRITISH LIBRARY. *Exchange Formats*. [S. l.: S. n.], 2003. Disponível em:

<<http://www.bl.uk/services/bibliographic/exchange.html>>. Acesso em: 17 set. 2005.

CAMPELLO, B. S. *Introdução ao controle bibliográfico*. 2. ed. Brasília: Briquet de Lemos, 2006.

CERVO, A. L.; BERVIAN, P. A. *Metodologia científica*. 5. ed. São Paulo: Prentice Hall, 2003.

CHAN, L. M. *Cataloging and classification: an introduction*. 3. Ed. Lanham, Maryland: The Scarecrow Press, 2007.

CHARTIER, R. *A ordem dos livros: leitores, autores e bibliotecas na Europa entre os séculos XIV e XVIII*. 2. ed. Tradução de Mary Del Priore. Brasília: Universidade de Brasília, 1998. (Coleção Tempos).

CHEN, P. *Modelagem de dados: a abordagem Entidade-Relacionamento para projeto lógico*. São Paulo: McGraw-Hill, MAKRON Books, 1990.

COAD, P. YOURDON, E. *Análise baseada em objetos*. 2. ed. Rio de Janeiro: Campus, 1991.

COSTA, M. A. F. da; COSTA, M. de F. B. da. *Metodologia da pesquisa: conceitos e técnicas*. Rio de Janeiro: Interciência, 2001.

DATE, C. J. *Introdução aos sistemas de bancos de dados*. Rio de Janeiro: Campus, 2004.

DEMPSEY, L.; HEERY, R. A review of metadata: a survey of current resource description formats. 1997. Disponível em: <http://www.ukoln.ac.uk/metadata/desire/overview/rev_ti.htm>. Acesso em: 30 ago. 2009.

DUVAL, E. et al. Metadata Principles and Practicalities. D-Lib Magazine, v. 8, n. 4, 2002. Disponível em: <<http://www.dlib.org/dlib/april02/weibel/04weibel.html>>. Acesso em: 30 ago. 2009.

FERREIRA, A. B. de H. *Novo dicionário Aurélio da língua portuguesa*. Curitiba: Positivo, 2004.

FERREIRA, A. B. de H. *Novo dicionário eletrônico Aurélio versão 5.0*. 2004. Corresponde ao Novo Dicionário Aurélio da Língua Portuguesa 3. ed., 1ª impressão, Editora Positivo.

FERREIRA, M. M. (Trad. e Adap.). *MARC 21: formato condensado para dados bibliográficos*. 2.ed. Marília: Universidade Estadual Paulista, 2002. v. 1.

FURRIE, B. *Understanding MARC Bibliographic: Machine – Readable Cataloging*. 2000. Disponível em: <<http://www.loc.gov/marc/umb/>>. Acesso em: 30 jul. 2007.

GILLILAND-SWETLAND, A. J. La definición de los metadatos. In: INTRODUCCIÓN a los metadatos: vías a la información digital. [S. l.]: GETTY, 1999. p. 1-9.

GÓMEZ, P. A. P. RDA: el alcance internacional del nuevo Código de Catalogación. *TransInformação*, Campinas, v. 19, n.3, p. 219-226, set./dez., 2007.

GRÁCIO, J. C. A. *Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade*. 2002. 127 f. Dissertação (Mestrado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002.

HILLMANN, D. *Using Dublin Core*. 2005. Disponível em: <<http://dublincore.org/documents/2005/11/07/usageguide/>>. Acesso em: 26 jul. 2009.

IFLA. *DECLARAÇÃO DOS PRINCÍPIOS INTERNACIONAIS DE CATALOGAÇÃO*. 2009. Disponível em: <http://archive.ifla.org/VII/s13/icp/ICP-2009_pt.pdf>. Acesso em: 27 jan. 2010.

IFLA. *Functional Requirements for Bibliographic Records: final report*. 1998. V. 19, UBCIM Publications – New Series. München: K.G. Saur, 1998. Disponível em: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>. Acesso em: 10 fev. 2006.

IFLA. *ISBD(M): Descrição Bibliográfica Internacional Normalizada para as Publicações Monográficas (International Standard Bibliographic Description for Monographic Publications)*. Lisboa: Biblioteca Nacional, 2005. Disponível em: <www.ifla.org/VI/3/nd1/isbdM-pt.pdf>. Acesso em 25 fev. 2006.

KÖCHE, J. C. *Fundamentos da metodologia Científica: teoria da ciência e iniciação à pesquisa*. 20. ed. Petrópolis: Vozes, 2002.

LE COADIC, Y. F. *A Ciência da Informação*. Brasília: Briquet de Lemos, 2004.

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 and XML design considerations*. [S. l.: S. n.], 2002. Disponível em: <<http://www.loc.gov/standards/marcxml/marcxml-design.html>>. Acesso em: 10 abr. 2005.

LITTON, G. *O livro e sua história*. Tradução de Maria Elvira Strong. São Paulo: McGraw-Hill do Brasil, 1975.

LYONS, M. Os novos leitores no século XIX: mulheres, crianças, operários. In: CAVALLO, G.; CHARTIER, R. (ORG.) *Leitura no mundo ocidental*. São Paulo: Ática, 1999. v.2, p. 135-163.

MACHADO, A. M. N. Informação e controle bibliográfico: um olhar sobre a cibernética. São Paulo: Unesp, 2003.

MARCONDES, C. H. Representação e economia da informação. *CiOnline*, Brasília, v.30, n.1, p. 61-70, 2001. Disponível em: <<http://www.scielo.br/pdf/ci/v30n1/a08v30n1.pdf>>. Acesso em: 06 fev. 2009.

MARTINS, W. *A palavra escrita: história do livro, da imprensa e da biblioteca*. 3. ed. il. rev. e at. São Paulo: Ática, 2001.

MÉNDEZ RODRÍGUEZ, E. *Metadados y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002. 429 p.

MEY, E. S. A.; SILVEIRA, N. C. *Catálogo no plural*. Brasília: Briquet de Lemos / Livros, 2009.

MILLER, P. *Metadata for the masses*. [S. l.: S. n.], 1996. Disponível em: <<http://www.ariadne.ac.uk/issue5/metadata-masses/>>. Acesso em: 16 abril 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 jun. 2004.

MORA, J. F. *Dicionário de filosofia*. São Paulo: Martins Fontes, 1996.

MORENO, F. P.; ARELLANO, M. Á. M. Requisitos Funcionais para Registros Bibliográficos - FRBR: uma apresentação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v.3, n. 1, p. 20-38, jul./dez. 2005.

NISO. Understanding Metadata. Bethesda: NISO Press, 2004. Disponível em: <<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>>. Acesso em: 25 jan. 2006.

ORTEGA, Cristina Dotta. Os registros de informação dos sistemas documentários: uma discussão no âmbito da Representação Descritiva. 250 f. Tese (Doutorado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo. São Paulo, 2009.

ORTIZ-REPISO JIMÉNEZ, V. Qué enseñamos después del MARC? *Organizacion del Conocimiento en Sistemas de Información y Documentación*, Zaragoza, v. 03, p. 217-225, 1999.

PEREIRA, A. M., SANTOS, P. L. V. A da C. O uso estratégico das tecnologias em catalogação. *Cadernos da F.F.C.*, Marília, v. 7, n. 1/2, p. 121- 131, 1998.

ROSETTO, M. *Metadados e formatos de metadados em sistemas de informação: caracterização e definição*. 2003, 95 f. Dissertação (Mestrado em Ciências da Comunicação)– Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003.

SANTOS, G. M. G. dos. *Dicionário júnior da língua portuguesa*. 2. ed. São Paulo: FTD. 2001.

SANTOS, P. L. V. A. da C., CORRÊA, R. M. R. Catalogação: trajetória para um código internacional. Niterói: INTERTEXTO, 2009.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SENSO, J. A.; ROSA PIÑERO, A. de la. El concepto de metadato. Algo más que descripción de recursos electrónicos. *Ciência da Informação*, Brasília, v. 32, n. 2, p. 95-106, maio/ago. 2003.

SILBERSCHATZ, A.; KORTH, H.F.; SUDARSHAN, S. *Sistema de banco de dados*. Rio de Janeiro: Elsevier, Campus, 2006.

SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. *Ciência da Informação*, Brasília, v.29, n.1, p.93-102, jan./abr. 2000. Disponível em: <<http://www.ibict.br/cionline/290100/29010010.pdf>>. Acesso em: 06 jul. 2000.

SOUZA, T. B. de, CATARINO, M. E., SANTOS, P. C. dos. Metadados: catalogando dados na Internet. *Transformação*, Campinas, v. 9, n. 2, maio/ago. 1997. Disponível em: <<http://www.puccamp.br/~biblio/tbsouza92.html>>. Acesso em: 12 jul. 2000.

TAKAHASHI, T. (Org.). *Sociedade da Informação no Brasil: o livro verde*. Brasília: Ministério da Ciência e Tecnologia, 2000.

WITTIMANN, R. Existe uma revolução da leitura no final do século XVIII? In: CAVALLO, G.; CHARTIER, R. (ORG.) *Leitura no mundo ocidental*. São Paulo: Ática, 1999. v.2, p. 135-163.

WOODLEY, M. S.; CLEMENT, G.; WINN, P. *DCMI Glossary*. [S. l.: S. n.], 2005. Disponível em: <<http://dublincore.org/documents/usageguide/glossary.shtml>>. Acesso em: 26 jul. 2009.

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