

XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB 2013)
GT 8: Informação e Tecnologia

Comunicação Oral

MODELO PARA A DESCRIÇÃO BIBLIOGRÁFICA E INTEROPERABILIDADE SEMÂNTICA

Fabiano Ferreira de Castro – UFS
Plácida L.V. A. C. Santos – UNESP

Resumo

O tratamento dado ao recurso informacional requer uma descrição de forma e de conteúdo legível por máquinas, com resultados compreensíveis aos humanos, e, que ainda, atenda aos requisitos de interoperabilidade entre os ambientes informacionais. Nesse sentido, a investigação propõe a reflexão e um olhar pautados numa proposta de desenvolvimento de um modelo para a descrição bibliográfica semântica, que garanta a interoperabilidade em ambientes informacionais digitais. A metodologia consiste na análise exploratória e descritiva do tema, verificando-se que as metodologias da Catalogação Descritiva e uma ontologia para a descrição bibliográfica, explicitada nas regras e códigos de catalogação, e nos padrões de metadados redesenham a construção de novos ambientes informacionais digitais melhor estruturados, para a recuperação da informação e no estabelecimento efetivo da interoperabilidade.

Palavras-chave: Interoperabilidade. Padrões de metadados. Modelagem de ambientes informacionais digitais. Informação e Tecnologia.

MODEL FOR THE SEMANTIC BIBLIOGRAPHIC DESCRIPTION AND INTEROPERABILITY

Abstract

The treatment of the information resource requires a description of form and content machine readable, understandable results to humans, and which also meets the requirements of interoperability between information environments. In this sense, the research proposes a reflection and a look lined a proposal to develop a model for the bibliographic description semantics, which ensures interoperability in digital informational environments. The methodology consists in analysis exploratory and descriptive of the subject, it appears that the methodologies of Descriptive Cataloging and ontology for bibliographic description, explained the rules and codes of cataloging and metadata standards to redesign the construction of new digital informational environments better structured for information retrieval and effective establishment of interoperability.

Keywords: Interoperability. Metadata patterns. Modeling of digital informational environments. Information and Technology.

1 INTRODUÇÃO

A preocupação da comunidade biblioteconômica advinda com as novas regras para a descrição bibliográfica aplicadas no ambiente digital levam os profissionais a repensarem o

seu fazer profissional, na tentativa de caminharem em consonância com as novas transformações trazidas no bojo da recontextualização tecnológica das bibliotecas.

Pode-se observar contemporaneamente um cenário marcado por grandes transformações tecnológicas e pelo surgimento e criação das mais variadas técnicas para as questões de representação, de descrição, de produção, de organização, de transmissão e de uso das informações em ambientes informacionais digitais.

O ambiente *Web*, por possuir informações heterogêneas e efêmeras, pode ser caracterizado como caótico, onde profissionais de diversas áreas buscam soluções para um tratamento mais significativo dos conteúdos que ali estão armazenados e disponibilizados para um uso mais efetivo. Idealizada por Tim Berners-Lee e liderada pelo *World Wide Web Consortium* (W3C) - a *Web Semântica* é um projeto que visa a solucionar e a minimizar os problemas de recuperação da informação na *Web* por meio do acesso automatizado aos recursos informacionais. Para tanto, pretende implantar uma estrutura tecnológica e instaurar uma maior representação do conhecimento na rede com o estabelecimento semântico das informações contidas nas unidades e assim, possibilitar técnicas mais eficazes de recuperação (BERNERS-LEE; HENDLER; LASSILA, 2001, tradução nossa).

Durante as últimas décadas observou-se que mais e mais informações têm sido publicadas, armazenadas e disponibilizadas na *Web*. No entanto, somente uma parte dessas informações e em ambientes específicos, tais como as bibliotecas digitais, por exemplo, possuem um padrão de metadados para a descrição bibliográfica e que possibilite a padronização dos recursos informacionais para atender aos requisitos da interoperabilidade.

As bibliotecas digitais apresentam-se e atuam como um segmento na Internet onde procuram desenvolver e criar métodos e técnicas para a padronização dos recursos informacionais. Mas garantir/assegurar a interoperabilidade entre “ilhas” de informações bem estruturadas e padronizadas entre formatos bibliográficos distintos encontra-se numa questão de investigação e preocupação pela comunidade científica. (CASTRO; SANTOS, 2010).

Novas tecnologias baseadas em pesquisas na temática *Web Semântica* desenvolvem-se com o intuito de tentar agregar semântica à descrição bibliográfica para um tratamento mais efetivo dos conteúdos digitais. A *Web Semântica* oferece algumas soluções para diferentes problemas criando uma nova visão para o armazenamento e o processamento dos dados. Algumas dessas soluções poderiam ser implementadas para potencializar os resultados das buscas no âmbito das bibliotecas digitais.

A preocupação da comunidade biblioteconômica advinda com as novas regras para a descrição bibliográfica aplicadas no ambiente digital, levam os profissionais a repensarem o

seu fazer profissional, na tentativa de caminharem em consonância com as novas transformações trazidas no bojo da recontextualização tecnológica das bibliotecas.

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

2 A CATALOGAÇÃO DESCRITIVA E SUA REPERCUSSÃO NOS AMBIENTES INFORMACIONAIS DIGITAIS

No cenário atual da comunidade da Biblioteconomia e da Ciência da Informação, há uma preocupação e um reconhecimento crescente de que haverá a necessidade de um sucessor do formato bibliográfico MARC 21 (*Machine Readable Cataloging*), devido às novas transformações no domínio bibliográfico, permeada pelo uso intensivo das TIC (Tecnologias da Informação e Comunicação). Segundo Coyle (2011) tais discussões tendem a se concentrar principalmente nas questões estruturais, o novo formato será XML (*eXtensible Markup Language*), ele fará uso de RDF (*Resource Description Framework*) e padrões de dados ligados¹?

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

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

¹ O termo dados ligados (*Linked Data*) refere-se ao conjunto das melhores práticas para a publicação e a interligação de dados estruturados na *Web*. Essas melhores práticas foram introduzidas por Tim Berners-Lee em sua nota *Web Architecture* e tornaram-se conhecidas como princípios de dados ligados. (HEATH; BIZER, 2011, tradução nossa).

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

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

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

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

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

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

Coyle (2012) reforça que o resultado é que agora tem-se uma série de “silos” RDF, expressões dos dados em RDF que faltam a capacidade de cruzamento dos dados ligados,

porque eles são obrigados a especificar estruturas de dados e pouco se ganha em termos de dados ligados sob a óptica bibliográfica. Não somente os esquemas RDF são incompatíveis com outros, ninguém será ligado a dados bibliográficos de comunidades fora de bibliotecas que publicarem seus dados na *Web*. O que significa a não ligação dos dados à Amazon, à *Wikipedia*, para citações em documentos.

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

- Definir "super-elementos" que estejam acima dos registros de formatos e que não estão vinculados pelas restrições dos registros definidos em RDF. Neste caso, haveria um "lugar de publicação" geral que seria um "super" elemento correspondente a todos os lugares de publicação nos vários registros e seriam subordinados a um conceito geral de "lugar" utilizado amplamente. Para implementar a ligação, cada elemento de registro seria extrapolado para os seus "super-elementos".
- Definir primeiro os elementos de dados fora de qualquer formato de registro específico, e então usá-los em esquemas de registros. Neste caso haveria apenas uma instância do "lugar de publicação" e que seria utilizado nos vários registros bibliográficos, sempre que um elemento é necessário. O intercâmbio desses registros seria possível com a ligação dos dados de seus elementos componentes, e iria interagir com outros dados bibliográficos na *Web* utilizando os elementos RDF definidos e seus relacionamentos.

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

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

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

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

3 MODELAGEM ESTRUTURAL DE DADOS PARA CONSTRUÇÃO DE AMBIENTES DIGITAIS

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

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

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

Qualquer mudança ou mapeamento que foi contratado para criar dados interoperáveis produziria um menor denominador comum (os dados mais simples e menos granular), e uma vez interoperáveis, não seria possível sua recuperação na íntegra, devido sua perda. Dados com menos estrutura e menos granularidade poderiam ser mais fáceis e mais baratos para aplicá-los e ter o potencial mais simples para as comunidades envolvidas. (YEE, 2009, p. 59, tradução nossa).

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

ao gênero do criador, onde o catalogador poderia se deparar com uma situação, que não saberia se o definiria como masculino ou feminino”. (YEE, 2009, p. 59, tradução nossa).

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

Na representação do campo 100 (autor pessoal), do padrão de metadados MARC 21, por exemplo, a forma autorizada para descrever o conteúdo é definida da seguinte maneira:

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

Nesse caso, o subcampo *\$d* (Datas associadas ao nome NR²), 1838 corresponde à data de nascimento, enquanto 1918 à data de falecimento do autor.

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

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

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

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

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

² (NR) Não Repetido, significa que só pode existir apenas um subcampo com apenas um código associado ao nome do responsável pela obra.

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

Castro (2012) define estrutura de dados, a camada intangível de instanciamento dos dados bibliográficos modelados para a representação e a descrição, tal como os formatos e/ou padrões de metadados, de forma a permitir a interoperabilidade dos ambientes informacionais digitais, por agentes humanos e não humanos, garantindo interfaces mais acessíveis aos usuários para posterior recuperação, uso e (re) uso dos recursos informacionais.

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

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

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

Para Castro (2012), os bibliotecários que se acostumarem com os Requisitos Funcionais para Registros Bibliográficos - FRBR - provavelmente não encontrarão muita dificuldade na transição do modelo conceitual para o modelo RDF. É importante nesse

momento destacar as principais terminologias empregadas pelos FRBR, RDF e RDFS, de acordo com o Quadro 1.

Quadro 1: Diferenças terminológicas nos modelos de dados.

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

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

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

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

A dificuldade em qualquer exercício de modelagem de dados, sobretudo, no domínio bibliográfico, encontra-se em decidir o que tratar como entidade ou classe e o que tratar como um atributo ou propriedade. FRBR decidiu criar uma classe chamada expressão para tratar qualquer mudança no conteúdo de uma obra. (CASTRO, 2012). Os FRBR em comparação ao

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

FRBR em RDF³ acrescenta apenas três classes. Duas delas (“*Endeavor*” e “Entidade Responsável”) são super conjuntos de classes FRBR. *Endeavor* é uma generalização que pode ser relacionada à obra, expressão, ou manifestação, ou seja, uma classe cujos membros são qualquer um dos produtos da atividade artística ou criativa. “Entidade Responsável” é um termo mais geral que pode se relacionar a uma entidade coletiva ou a uma pessoa. Essas classes especificam mais claramente informações sobre o conteúdo intelectual de um recurso, sem necessitar fornecer informações adicionais. “A terceira classe que é adicionada é o assunto. Todas as três incluem alguma instância do assunto em seus esquemas. FRBR trata claramente assunto como um relacionamento”. (DAVIS; NEWMAN, 2005; COYLE, 2012).

A condição dos ambientes informacionais digitais interoperar seus dados é a modelagem conceitual, esta definida e codificada por uma série de requisitos funcionais estabelecidos pelas arquiteturas de metadados, regras e esquemas de descrição bibliográfica e ontologias, que fornecerá ambientes melhor estruturados, na garantia de recuperação de informações mais efetiva aos usuários (humanos e não humanos).

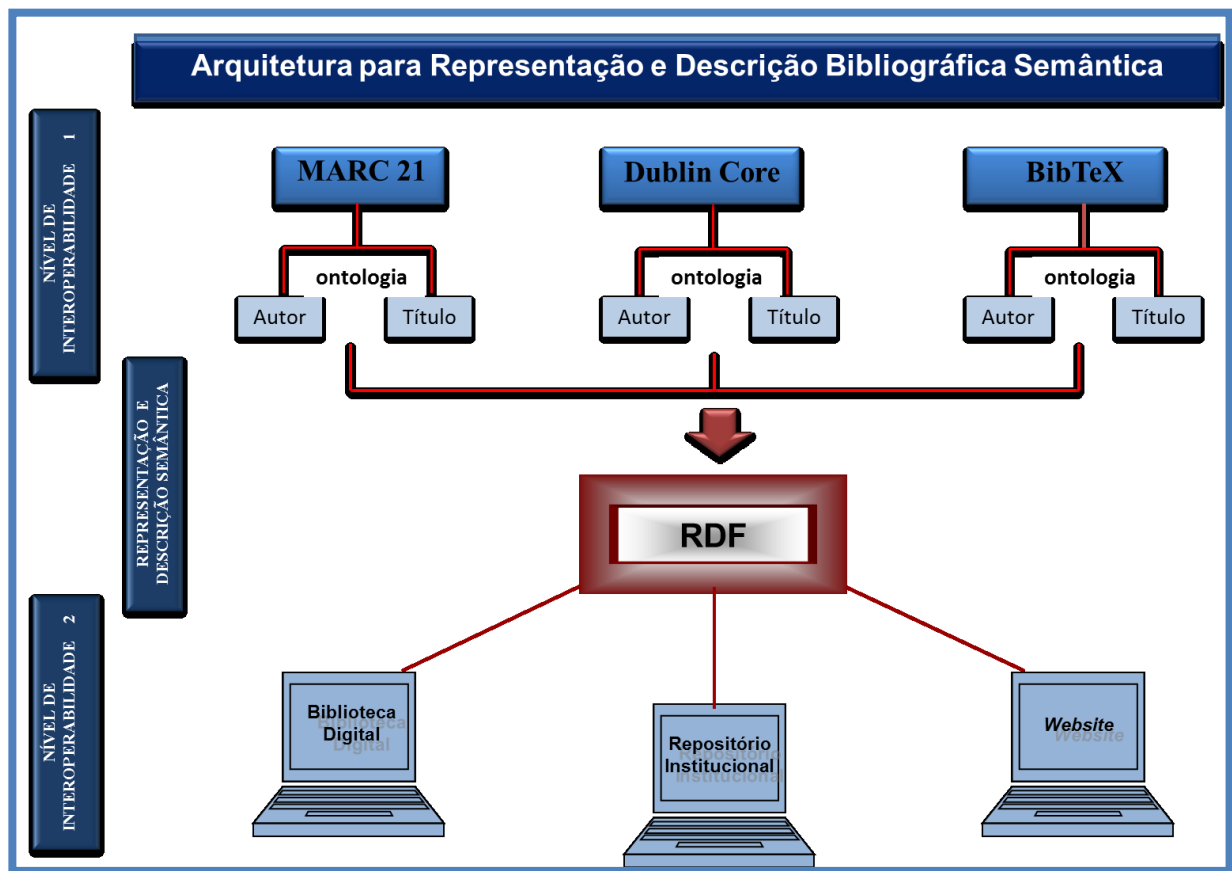
4 APRESENTAÇÃO DE UM MODELO PARA A DESCRIÇÃO BIBLIOGRÁFICA SEMÂNTICA EM AMBIENTE DIGITAL: À GUIA DE RESULTADOS

Percebe-se que a palavra de ordem na constituição de ambientes informacionais digitais, sobretudo, no âmbito das bibliotecas digitais é a interoperabilidade. Para tanto, para que a interoperabilidade aconteça de maneira eficiente, faz-se necessário um olhar mais acurado nas camadas estruturantes e delineadoras dos catálogos bibliográficos digitais, ou seja, na representação e na descrição dos dados, a fim de potencializar as formas de busca e recuperação das informações.

Com base na literatura científica apresentada sobre a representação, a descrição, a modelagem e a interoperabilidade de recursos bibliográficos e pautados nas principais definições, arquiteturas e tecnologias, e como resultados dessa investigação, Castro (2012) apresenta uma proposta de um modelo/esquema para a representação e a descrição bibliográfica semântica de um recurso informacional, contempladas num único ambiente, com vistas a possibilitar a interoperabilidade semântica entre padrões de metadados e ambientes informacionais heterogêneos, de acordo com a figura 1.

³ As expressões do núcleo de conceitos e vocabulários FRBR em RDF podem ser melhor visualizados no *site*. Disponível em: <<http://vocab.org/frbr/core.html>>. Acesso em: 20 jan. 2013.

FIGURA 1: Modelo para a Representação e a Descrição Bibliográfica semântica e níveis de interoperabilidade.



Dublin Core e BibTeX), descrevendo alguns metadados em comum, tais como, autor e título (considerados elementos de calção de um recurso bibliográfico). Na fase da definição dos metadados precisa-se considerar para estruturar os elementos as ontologias, ou seja, o delineamento conceitual de cada atributo e os modelos conceituais de dados, tais como os FRBR e FRAD.

No padrão de metadados e/ou formato de intercâmbio MARC 21, por exemplo, as ontologias são determinadas através das regras e esquemas de descrição bibliográfica (AACR2) e também, do conceito fornecido pelo próprio MARC 21.

Quadro 2: Explicitação de ontologia para os elementos autor e título em AACR2.

Ontologia para o esquema descrição bibliográfica em AACR2	
Área	Ontologia
Autor	A pessoa a quem cabe à responsabilidade principal pela criação do conteúdo intelectual ou artístico de uma obra.

Título	Palavra, frase, caractere, ou grupo de caracteres, que normalmente aparecem num item, dando nome a estes ou à obra nele contida.
---------------	--

Fonte: Elaborado pelo autor a partir do AACR2 (Glossário D).

A definição dos elementos de metadados pela ontologia de descrição bibliográfica é considerada uma das principais partes da interoperabilidade por poder fornecer informações significativas para que sejam compreensíveis por agentes não humanos (computador/máquina).

Quadro 3: Explicitação de ontologia para elementos autor e título em MARC 21.

Ontologia para o formato MARC 21	
Etiqueta	Ontologia
100 (Autor)	Nome pessoal usado como uma entrada principal em um registro bibliográfico. A entrada principal é atribuída de acordo com as várias regras de catalogação, usualmente para o pessoal responsável pela criação intelectual de uma obra.
245 (Título)	Título e indicação de responsabilidade são áreas da descrição bibliográfica de uma obra. A indicação do título consiste da propriedade do título e também pode conter a designação geral do material (Meio), complemento do título, outras informações sobre o título, a complementação do título da página de rosto, e a indicação de responsabilidade. A propriedade do título inclui o título abreviado (<i>short title</i>) e o título alternativo (<i>alternative title</i>), a indicação numérica de uma parte/seção e o nome de uma parte/seção.

Fonte: Elaborado pelo autor a partir da *Library of Congress* (2011).

Percebem-se conceitos semânticos na ontologia de descrição bibliográfica AACR2 e MARC 21, pautados na lógica de descrição estabelecidos no processo da Catalogação Descritiva, uma vez que a base de conhecimento do formato MARC 21 é o AACR2. No RDA, por exemplo, não é explicitado o conceito Autor, o mesmo é chamado por *Creator* (Criador), linguagem que está em consonância com o padrão de metadados Dublin Core, desenhado para a localização de recursos na *Web*.

Vale destacar nesse momento que a arquitetura supracitada apresenta ontologias para apenas dois elementos da descrição bibliográfica na estrutura MARC 21 (autor e título), uma

vez que é necessário o desenvolvimento e a criação de ontologias para determinar a semântica de todos os metadados do padrão MARC 21, pois quanto mais complexa e completa for a estruturação, haverá uma compreensão mais intuitiva e efetiva pela máquina promovendo a interoperabilidade. (CASTRO, 2012).

Ainda na fase de confecção dos metadados, a atribuição das ontologias para a descrição bibliográfica não é condição suficiente para determinar a interoperabilidade. O catalogador além de conhecer a estrutura MARC 21 deverá também dominar e refletir o uso das regras de catalogação (AACR2) para saber aplicá-las no preenchimento do conteúdo dos metadados (valores dos elementos).

Para a descrição bibliográfica de um recurso bibliográfico (livro, por exemplo) na estrutura MARC 21, o catalogador deverá seguir as regras do capítulo 2 do código de catalogação (AACR2) que norteiam a construção padronizada e unívoca de como deverão ser representados e apresentados os dados aos usuários. Os metadados título e autor devem seguir as seguintes instruções, conforme o Quadro 4.

Quadro 4: Regras de conteúdo definidas por AACR2 para elementos autor e título em MARC 21.

Regras de construção padronizada dos valores dos metadados em MARC 21, segundo o AACR2.	
Regra	Definição
2.1B Título	Transcreva o título principal de acordo com as instruções de 1.1 B.
1.1 B1.	Transcreva o título principal exatamente no que respeita à redação, ordem e grafia, mas não necessariamente quanto à pontuação e ao uso de maiúsculas. Use os acentos e outros sinais diacríticos que estiverem na fonte principal de informação (veja também 1.0 G). Use maiúsculas de acordo com o apêndice A.

2.1 F.	Indicação responsabilidade
2.1 F1.	Transcreva indicações de responsabilidade relativas às pessoas ou entidades de acordo com as instruções de 1.1 F.
1.1 F1.	Transcreva indicações de responsabilidade que figurem com destaque no item na forma em que eles aparecem. Se uma indicação de responsabilidade for extraída de uma fonte que não seja a fonte principal de informação, coloque-a entre colchetes.

Fonte: Elaborado pelo autor a partir do AACR2 (2005).

É através do uso adequado das regras e dos esquemas que as representações terão maior consistência na camada intangível do sistema propiciando formas de apresentação dos recursos bibliográficos de forma mais efetiva.

A meta no nível de interoperabilidade 1, de acordo com a figura 1, é a possibilidade de integração e comunicação entre padrões de metadados distintos, para tanto só será possível a partir da determinação de ontologias de descrição bibliográfica explicitadas nas regras e códigos de catalogação e em formatos de metadados potencializando os relacionamentos bibliográficos semânticos e estabelecendo a interoperabilidade.

A camada RDF terá o objetivo de fazer o armazenamento e a representação dos dados, localizada na parte central da arquitetura, com as descrições bibliográficas semânticas. Independente da estrutura original do padrão permitirá que os dados em MARC 21, Dublin Core, BibTeX ou qualquer outro padrão, seja convertido em RDF e podendo ser reaproveitados na construção *Web Semântica*.

Já na parte inferior da arquitetura é possível verificar a presença de ambientes informacionais heterogêneos, que podem ser interpretados tais como, Biblioteca Digital, Repositório Institucional, *Webside* etc., definidos no modelo, sendo que a representação e a descrição dos recursos informacionais, podem estar ou não contemplados num mesmo padrão. O nível de interoperabilidade 2 objetiva a possibilidade de interoperabilidade semântica entre múltiplos ambientes e sistemas informacionais digitais e o armazenamento das descrições bibliográficas semânticas numa única plataforma/interface.

Ressalta-se, que quanto mais completa e complexa for uma descrição, melhores possibilidades nos níveis de interoperabilidade semântica haverá, com a possibilidade de uso, de preservação e de (re) uso das informações de forma mais efetiva pelo usuário final, multidimensionando as formas de acesso aos recursos informacionais.

Portanto, acredita-se tal modelo/esquema de arquitetura proposto nessa investigação, pode possibilitar o compartilhamento entre padrões de metadados e ambientes e sistemas informacionais distintos, trabalhando numa filosofia de colaboração entre os recursos informacionais disponíveis e as tecnologias que estão abarcadas na sua construção, no estabelecimento da interoperabilidade, na otimização dos relacionamentos bibliográficos e ampliados para a construção padronizada de recursos na *Web*.

5 CONSIDERAÇÕES FINAIS

Verifica-se, no panorama atual da Catalogação Descritiva, que os modelos conceituais propostos pelos Requisitos Funcionais para Registros Bibliográficos (FRBR) e as bases ontológicas dos esquemas de metadados podem propiciar a clareza nos relacionamentos entre registros bibliográficos.

Vale dizer que o modelo FRBR facilita o desenho de um modelo conceitual, em consonância com as ontologias e os metadados, não somente por sua riqueza conceitual e estrutural, mas porque se constitui num marco de referência para a compreensão dos relacionamentos bibliográficos e na modelagem dos catálogos digitais.

O modelo RDF subjacente às tecnologias da *Web Semântica* é frequentemente descrito como o futuro de metadados estruturados. A sua adoção em bibliotecas tem sido lenta, justificado pelo fato de que as diferenças fundamentais estão na abordagem de modelagem que leva RDF, representando uma arquitetura "*bottom up*" onde as descrições são distribuídas e pode ser feita sob qualquer característica considerada necessária, enquanto a abordagem "cêntrica" de registros bibliográficos das bibliotecas tende a ser mais "*top down*" confiando nas funções pré-definidas determinadas por elas. (LEE; JACOB, 2011; YEE, 2009; RILEY, 2010; COYLE, 2012).

MARC 21 fornece o protocolo pelo qual os computadores realizam o intercâmbio, o uso e a interpretação da informação bibliográfica, ligando elementos de dados para formar a base da maioria dos catálogos de bibliotecas utilizada hoje. A *Web Semântica* em muitos aspectos será baseada na utilização desse tipo de ligação de dados, mas talvez em um catálogo ou banco de dados muito maior e globalizado.

O uso dos modelos conceituais de dados, das arquiteturas de metadados e das ontologias redesenham os novos ambientes informacionais digitais, definindo conceitualmente os elementos da descrição bibliográfica a serem representados pelo catalogador, proporcionando interfaces de buscas mais compreensíveis aos usuários e no estabelecimento efetivo da interoperabilidade.

A adoção do padrão de metadados não é condição suficiente para atender aos requisitos de interoperabilidade. Faz-se necessário a utilização correta das regras e/ou esquemas de descrição, pautados numa lógica descritiva, para a definição e construção padronizada dos metadados gerando representações dos recursos bibliográficos consistentes e unívocas, alcançando dessa forma, as potencialidades da interoperabilidade em ambientes informacionais digitais. (CASTRO, 2012).

Quanto ao “futuro” do MARC 21, sabe-se que há iniciativas internacionais de construção de conversores, ferramentas e instrumentos tecnológicos que podem potencializar a transformação dos dados bibliográficos, numa estrutura RDF, mas ainda nenhuma oficializada pelos organismos internacionais de catalogação. Essa investigação não tem a pretensão de desenvolver o que poderia ser chamado de MARC RDF, mas refletir sobre um caminho inicial com orientações teóricas e metodológicas, para a determinação e a construção de ambientes informacionais melhor estruturados, com descrições bibliográficas semânticas explicitadas, propiciando a interoperabilidade em ambientes informacionais digitais.

Acredita-se que o modelo para a representação e a descrição bibliográfica semântica e os níveis de interoperabilidade desenvolvida nesse trabalho, propicia a modelagem estrutural dos ambientes informacionais digitais atualmente, a partir da heterogeneidade de esquemas de metadados, tornando-se, dessa forma, uma estrutura única a ser adotada por vários catalogadores, ampliando seu escopo para a construção padronizada de recursos na *Web*. Além disso, a adoção desse modelo configura o estabelecimento efetivo da interoperabilidade, principalmente com os impactos tecnológicos na Catalogação Descritiva e o desenvolvimento de ambientes informacionais digitais interoperáveis.

REFERÊNCIAS

AACR2. *Anglo American Cataloging Rules*. Preparado sob a direção de The Joint Steering Committee for Revision of AACR; trad. Federação Brasileira de Associações de Bibliotecários, Cientistas da Informação e Instituições (FEBAB). 2. ed., rev. 2002. São Paulo: FEBAB, 2005.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, New York, May, 2001. Disponível em: <<http://www.sciam.com/2001/0501issue/0501berners-lee.html>>. Acesso em: 28 jun. 2012.

CASTRO, Fabiano F. de. Elementos de interoperabilidade na catalogação descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais. 2012. 202 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.

CASTRO, Fabiano. F. de; SANTOS, Plácida L. V. A. C. Catalogação e metadados: interlocuções nos ambientes informacionais digitais. In: ARELLANO, Filiberto Felipe Martínez (Org.). *Memoria del IV encuentro de catalogación y metadatos*. México: Centro Universitario de Investigaciones Bibliotecológicas de la Univ. Nacional Autónoma de México, 2010. v. 4, p. 301-318.

COYLE, Karen. *Bibliographic framework: RDF and linked data*. Berkeley, United States, Jan. 2012. Disponível em: < <http://kcoyle.blogspot.com/2012/01/bibliographic-framework-rdf-and-linked.html>>. Acesso em: 20 jan. 2012.

DAVIS, Ian; NEWMAN, Richard. *Expression of core FRBR concepts in RDF*. 2005. Disponível em: < <http://vocab.org/frbr/core.html>>. Acesso em: 20 jan. 2012.

HEATH, Tom; BIZER, Christian. *Linked data: evolving the web into a global data space*. Morgan & Claypool, Vol. 1 no. 1, 2011. 136 p. (Synthesis Lectures on the Semantic Web: Theory and Technology). Disponível em: < <http://linkeddatabook.com/editions/1.0/#linkedData>>. Acesso em: 15 nov. 2011.

LEE, Seungmin; JACOB, Elin K. An integrated approach to metadata interoperability: construction of a conceptual structure between MARC and FRBR. *Library Resources & Technical Services*, Vol. 55, no. 1, p.17-32, Jan. 2011.

LIBRARY OF CONGRESS. *MARC 21 format for bibliographic data*. Library of Congress Network Development and MARC Standards Office, No. 13, Sep. 2011. Disponível em: <www.loc.gov/marc/bibliographic/ecbdhome.html>. Acesso em: 12 jan. 2012.

MORENO, F. P. *Requisitos funcionais para registros bibliográficos - FRBR: um estudo no catálogo da rede bibliodata*. 2006. 202 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2006.

RILEY, Jenn. *RDF for librarians*. DLP Brown Bag Series, p. 1-38, Sep. 2010. Disponível em: < <http://www.dlib.indiana.edu/education/brownbags/fall2010/rdf/rdf.pdf>>. Acesso em: 15 nov. 2011.

THOMALE, Jason. *Interpreting MARC: where's the bibliographic data?* *Code4Lib Journal*, Vol. 11, 2010.

YEE, Martha M. Can bibliographic data be put directly onto the semantic web? *Information Technology and Libraries*, Vol. 28, no. 2, p. 55-80, Jun. 2009.