

ELEMENTOS DE INTEROPERABILIDADE NA PERSPECTIVA DA CATALOGAÇÃO DESCRITIVA

*Fabiano Ferreira de Castro**
*Plácida L.V.A.C. Santos***

RESUMO

Atualmente, no ambiente informacional digital, onde vários tipos de recursos coexistem com formatos e padrões de metadados heterogêneos, muitos esforços têm sido feitos para alcançar a interoperabilidade, a fim de utilizar vários padrões de metadados e de reutilizar os registros de metadados, através do desenvolvimento de estratégias, que vão desde os simples mapeamentos entre elementos de metadados à modelagem estrutural complexa. O uso intensivo das Tecnologias da Informação e Comunicação (TIC) gera o desenvolvimento de uma gama de padrões de metadados, instrumentos fundamentais nos ambientes de manipulação de dados bibliográficos. Objetiva-se apontar em que medida as metodologias da Catalogação Descritiva, explicitadas nos códigos de catalogação (AACR2 e RDA) e nos padrões de metadados, favorecem o delineamento modelar de ambientes informacionais digitais melhor estruturados para a recuperação da informação. Para tanto, com uma metodologia de análise exploratória e descritiva do tema, com o estudo das principais produções científicas da área de Ciência da Informação e da Biblioteconomia, é possível identificar e compreender as principais mudanças tecnológicas no cerne da Catalogação Descritiva, no âmbito digital. Como resultado apresenta-se uma proposta de requisitos funcionais para a modelagem estrutural de ambientes informacionais digitais, que garantem a interoperabilidade no domínio bibliográfico. Acredita-se que a utilização dos metadados, das ontologias e dos fundamentos teóricos e metodológicos da Catalogação Descritiva, uma vez articuladas, podem redesenhar novas opções aos ambientes informacionais, seja na modelagem dos catálogos bibliográficos, ou ainda, na forma de representação dos recursos informacionais, garantindo a interoperabilidade.

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

* Doutor em Ciência da Informação pela Universidade Estadual Paulista Julio de Mesquita Filho, Campus Marília, Brasil. Professor do Departamento de Ciência da Informação da Universidade Federal de São Carlos, Brasil. Líder do Grupo de Pesquisas CNPq Estudos em Representação do Conhecimento e Tecnologia de Informação e Comunicação
E-mail: fabianofcastro.ufs@gmail.com.

** Doutora em Semiótica e Linguística Geral pela Universidade de São Paulo, Brasil. Livre-docente em Catalogação pela Universidade Estadual Paulista Julio de Mesquita Filho, Brasil. Docente permanente do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista Julio de Mesquita Filho, Brasil.
E-mail: placidasantos@gmail.com.

I INTRODUÇÃO

Contemporaneamente tem-se percebido um aumento desordenado e caótico na quantidade de informações produzidas e disponibilizadas em meio digital, requerendo uma mudança e um repensar nas formas de armazenamento, de representação, de descrição e de preservação dos recursos digitais em ambientes informacionais.

No tocante à evolução dos ambientes informacionais - as bibliotecas - têm merecido grande destaque, ao longo do século XX, no que se refere ao desenvolvimento e uso de tecnologias, principalmente as TIC, pois potencializaram seus serviços com o oferecimento de novos recursos de acesso, com a utilização de formatos de intercâmbio de dados na Catalogação Descritiva, no uso de ontologias, na

orientação para a modelagem de catálogos e no processo de busca e recuperação da informação.

É possível identificar que há uma tendência para a disponibilização de recursos e conteúdos informacionais em formato digital e a não utilização de padrões adequados para representar e descrever estes recursos. Elucida-se que o tratamento dado ao recurso informacional requer uma descrição de forma e de conteúdo legível por máquinas, com resultados compreensíveis aos humanos, e, que ainda, atenda aos requisitos de interoperabilidade entre os ambientes informacionais, apontados pela literatura e focados nesta investigação, enquanto problema da pesquisa.

Com os novos cenários da Catalogação Descritiva, marcados com a modelagem conceitual de catálogos com os Requisitos Funcionais para Registros Bibliográficos (FRBR), e a proposta do novo código de catalogação *Resource Description and Access* (RDA) e caracterizados por novas perspectivas tecnológicas, conhecidas hoje por *Web Semântica*, *Web 2.0*¹ e *Web 3.0*², a pesquisa tem por objetivo geral o estudo das ferramentas e dos métodos para o Tratamento Descritivo da Informação (TDI), em especial, para a representação e a descrição (forma e conteúdo) de recursos bibliográficos, no contexto digital.

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

Como resultados e reflexões destaca-se, em que medida os ambientes informacionais digitais precisam ser repensados a partir da modelagem conceitual, na camada intangível dos sistemas, definida na representação e na descrição dos dados bibliográficos, com alcance à interoperabilidade.

2 A DEFINIÇÃO DA INTEROPERABILIDADE NO DOMÍNIO BIBLIOGRÁFICO³

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

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

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

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

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

3 O termo domínio bibliográfico "[...] designa o campo de estudo voltado para o tratamento descritivo da informação (processo de representação do recurso) em bibliotecas". (ALVES; SANTOS, 2013, p. 16).

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

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

De acordo com Rodríguez:

A interoperabilidade supõe complementar uma série de estratégias humanas ligadas à descrição consistente e flexível da informação mediante metadados, com um conjunto de ferramentas, softwares e protocolos que permitam os distintos sistemas de informação (bibliotecas digitais) estabelecer mecanismos de conversão automática entre eles, e ainda, reconciliar suas diferenças e alcançar, ao menos, um acesso universal integrado aos recursos informacionais. (RODRÍGUEZ, 2002, p. 312, tradução nossa)

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

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

interdisciplinares subjacentes. (SAYÃO; MARCONDES, 2008).

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

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

- **Interoperabilidade técnica** – as considerações sobre os aspectos técnicos incluem assegurar envolvimento de um conjunto de organizações no contínuo desenvolvimento de padrões de comunicação, transporte, armazenamento e representação de informações, tais como são o Z39.50 e o *Search Retrieval Web Service* (SRW). Inclui também os esforços cooperativos para assegurar que padrões individuais evoluam em benefício da comunidade envolvida e para facilitar, onde for possível, convergência desses padrões, de forma que seja possível que os sistemas possam ter como base mais de um conjunto de padrões.
- **Interoperabilidade semântica** – está relacionada com o significado ou semântica das informações originadas de diferentes recursos e é solucionada pela adoção de ferramentas comuns ou/e mapeáveis de representação da informação, como esquemas de metadados, classificações, tesouros e mais recentemente, ontologias; um exemplo de questão endereçada por essa faceta da interoperabilidade pode ser o seguinte: o que significa “autor” para um recurso informacional? Será a mesma coisa que “criador” para um outro recurso? No nível semântico Marino (2001) acrescenta que existem dois subníveis: Epistemológico (relacionado ao significado dos elementos descritores do formato e das relações nele existentes) e Ontológico (relacionado ao uso de ontologias, vocabulários controlados e padrões de metadados para o estabelecimento dos significados dos dados representados). Como exemplo desse nível de interoperabilidade pode-se

destacar os metadados que têm o mesmo significado entre elementos de descrição do formato MARC 21.

- **Interoperabilidade política/humana** – independente das questões relacionadas à maneira pela qual a informação é descrita e disseminada, a decisão de tornar os recursos informacionais mais amplamente disponíveis e interoperáveis tem implicações para a organização, para as equipes envolvidas e para os usuários em termos comportamentais, de recursos e de treinamento. A ênfase dada por parte de alguns setores governamentais aos problemas de democratização do acesso, da exclusão digital e da federação de fontes de informação voltadas para a educação a distância, tem impacto nas políticas públicas para a área, e estão enquadrados nesse item.
- **Interoperabilidade intercomunitária** – enfoca a necessidade, cada vez mais urgente, impulsionada pela crescente interdisciplinaridade, principalmente nas áreas de pesquisa, de acesso a informações provenientes de um espectro amplo de fontes distribuídas por organizações, áreas de conhecimento e comunidades de natureza distintas. Geralmente exige o estabelecimento de fóruns para discussão e consenso em torno de práticas e procedimentos comuns.
- **Interoperabilidade legal** – considera as exigências e as implicações legais de tornar livremente disponíveis itens de informação.
- **Interoperabilidade internacional** – quando se atua em escala internacional é necessário contornar a diversidade de padrões e normas, os problemas de comunicação, as barreiras linguísticas, as diferenças no estilo de comunicação e na falta de uma fundamentação comum.

Vale destacar que essa investigação tem como interesse os requisitos de interoperabilidade nas esferas semântica, estrutural e sintática, onde defende-se que a estruturação e a modelagem de catálogos bibliográficos (banco de dados) ocorrem nessas camadas, para que a interoperabilidade entre ambientes informacionais digitais heterogêneos

ocorra de maneira efetiva e se estabeleçam os relacionamentos bibliográficos semânticos numa mesma interface.

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

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

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

Como podemos definir conceitualmente interoperabilidade no contexto dos metadados? O IEEE (2000) diz que a interoperabilidade corresponde à habilidade de dois ou mais sistemas ou componentes trocar informações e utilizá-las neste processo. Nilsson (2010) diz que a interoperabilidade de metadados consiste na capacidade de dois ou mais sistemas, ambientes informacionais ou componentes trocarem dados descritivos sobre recursos informacionais, e interpretá-los de uma maneira consistente, não ocorrendo falhas no intercâmbio para sua compreensão por máquinas.

De acordo com Sayão e Marcondes (2008),

Uma organização verdadeiramente interoperável é capaz de maximizar o valor e o potencial de reuso da informação que está sob o seu controle. É também capaz de intercambiar efetivamente estas informações com outras organizações igualmente interoperáveis, permitindo que novos conhecimentos possam ser gerados a partir da identificação de

relacionamentos entre conjuntos de dados previamente não relacionados. Na perspectiva do usuário, as interfaces devem apresentar para o usuário uma visão unificada em termos semânticos de diferentes recursos informacionais heterogêneos, ou seja: como nomeá-los, como referenciá-los, como utilizá-los em buscas, como acessá-los, como apresentá-los para o usuário. (SAYÃO; MARCONDES, 2008, P.136)

De acordo com Castro (2012), pode-se dizer que a interoperabilidade é a capacidade de dois ou mais sistemas, ambientes informacionais digitais e padrões de metadados heterogêneos, intercambiarem seus dados, a partir da codificação das regras e esquemas de descrição utilizadas pelo catalogador, na confecção padronizada e metodológica dos metadados, e na estruturação conceitual do ambiente informacional proporcionados pelas ontologias. Nesse sentido, Alves (2005, p. 110) diz que as ontologias funcionam como “[...] qualificadores dos metadados e não haverá estabelecimento da semântica sem a representação por metadados [...]”.

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

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

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

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

Neste cenário, pode-se destacar no domínio bibliográfico, a padronização do processo de catalogação que permite a construção metodológica dos metadados tem sido a questão desde as primeiras bibliotecas. Como os dados de uma biblioteca (como título, autor e assunto) geralmente estão abertos e disponíveis ao mundo, os principais problemas na interoperabilidade de metadados começam a surgir.

O desenvolvimento de um novo código de catalogação *Resource Description and Access* (RDA) aparece nesse contexto, como uma forma de auxiliar a interoperabilidade de metadados independente do ambiente informacional, seja ele uma biblioteca convencional ou digital, um centro de documentação, um repositório institucional, um arquivo etc. destacando a situação complexa com uma multiplicidade de padrões de metadados utilizados em nos ambientes informacionais que garantam a interoperabilidade.

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

Quando os metadados projetados de acordo com diferentes especificações e de diferentes domínios se encontram, ou seja, quando comunidades se envolvem para aumentar sua interação, dificuldades consideráveis no gerenciamento dos metadados tendem a aparecer, pois muitas vezes os dados se apresentam incompatíveis. O resultado é que os benefícios da interoperabilidade de metadados dentro de um padrão são perdidos quando são combinados, aumentam os custos de desenvolvimento e falhas na comunicação entre os sistemas, aplicando-se em muitos casos soluções não reutilizáveis.

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

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

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

3 ARQUITETURA DE METADADOS RDF, ESQUEMAS DE DESCRIÇÃO E MODELOS BIBLIOGRÁFICOS

Antes de iniciarmos as discussões sobre a arquitetura de metadados, os esquemas de descrição e os modelos bibliográficos elucidam-se alguns conceitos-chave para melhor compreensão das relações bibliográficas.

O *Resource Description and Access* (RDA) constitui-se de orientações para a descrição de recursos informacionais em ambientes digitais e baseia-se numa estrutura teórica e metodológica que define a forma e o conteúdo de um recurso. Baseia-se nos modelos conceituais *Functional Requirements for Bibliographic Records* (FRBR) e nos *Functional Requirements for Authority Data* (FRAD).

Os FRBR constituem-se num modelo conceitual para a modelagem de catálogos bibliográficos e fornecem uma nova perspectiva da estrutura e nos relacionamentos dos registros bibliográficos, pois relacionam os metadados presentes em um registro bibliográfico com as necessidades dos usuários e recomendam um nível básico de funcionalidade para esses registros. Os FRAD constituem-se num modelo conceitual para dados de autoridade e provê um quadro de referência, claramente definido e estruturado, para relacionar os dados que são registrados em registros de autoridade às necessidades dos usuários desses registros, além de auxiliar em uma avaliação do potencial para o compartilhamento internacional e uso de dados de autoridade em bibliotecas e em outros ambientes informacionais. (TAYLOR, 2007).

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

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

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

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

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

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

Da mesma forma, valores de vocabulários em *Simple Knowledge Organization System* (SKOS⁴)

⁴ SKOS is built upon and, and its main objective is to enable easy publication of controlled structured vocabularies for the. Família de linguagens formais criada para a representação de tesouros, esquemas de classificação, taxonomia, sistemas de cabeçalhos de assunto, ou qualquer outro tipo estruturado de vocabulário controlado. SKOS é construída sobre RDF e RDFS, e seu objetivo principal é permitir a fácil publicação de vocabulários controlados estruturados para a Web Semântica. Atualmente SKOS is currently developed within the framework. SKOS é desenvolvido pelo grupo

e um vocabulário RDF construído em OWL, permitem que os vocabulários sejam estendidos e usados de acordo com as intenções almejadas em RDA. O uso de SKOS permite também ao RDA ser mais bem integrado ao desenvolvimento geral de ontologias e organização do conhecimento que são importantes para melhorar as aplicações de recuperação da informação centradas nos usuários.

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

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

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

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

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

de trabalho e especificações do W3C. Maiores informações em: Disponível em: <<http://www.w3.org/TR/skos-reference/skos.html#semantic-relations>>. Acesso em: 20 jan. 2014.

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

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

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

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

Um aspecto significativo do RDA é o uso do modelo entidade-relacionamento FRBR. O uso do modelo entidade-relacionamento (ER) é um componente-chave da *Web Semântica*, mas é inteiramente novo nas regras de catalogação em bibliotecas. As revisões do RDA realizadas pelo JSC, ao longo de seu desenvolvimento, fizeram o comitê repensar as bases conceituais tradicionais estabelecidas pelas regras de catalogação. É importante destacar que o desenvolvimento dos elementos de RDA reflete o comprometimento do JSC junto aos princípios dos FRBR, embora a interação de RDA com os FRBR não tenha sido desenvolvida de tal forma

que traduziria diretamente a correspondência dos relacionamentos em RDF (HILLMANN et al., 2010). Nesse sentido, o grupo DCMI/RDA tem trabalhado no desenvolvimento e no registro de vocabulários para sanar as principais lacunas que possam existir na correspondência dos elementos RDA nas representações em RDF.

A integração entre RDA e FRBR significava que RDA em RDF seria necessário para interagir com um RDF compatível ao tratamento de FRBR, preferencialmente com um acordo oficial entre os seus criadores. No entanto, a IFLA foi incapaz de fornecer as entidades FRBR com os requisitos e identificadores necessários estruturados em RDF em tempo hábil. Além disso, o trabalho da IFLA sobre FRAD só recentemente chegou ao seu fim. Portanto, para realizar a criação de propriedades RDA em RDF foi necessário criar uma versão transacional de FRBR, preferencialmente no mesmo registro de RDA. Esta versão específica RDA de FRBR inclui “Família”, dos FRAD, e “Agente” do modelo orientado a objeto de FRBR, e inclui apenas o nível da entidade, com as entidades identificadas como classes e subclasses. Uma versão RDF de FRBR está oficialmente disponível a partir da IFLA, os relacionamentos entre as mesmas classes definidas na versão RDA de FRBR e uma versão IFLA dos FRBR será feita para indicar que estas são realmente as mesmas entidades. (HILLMANN et al., 2010).

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

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

De acordo com Hillmann et al. (2010), os estudos realizados por bibliotecários de música e catalogadores de audiovisuais mostram que a definição de entidades do Grupo 1 dos FRBR é altamente diferenciada quando se lida com diferentes formas de expressão criativa. Além

disso, há elementos de RDA que estão vinculados explicitamente às orientações de RDA para uma entidade FRBR, geralmente em virtude da inclusão da entidade em nome do elemento. Exemplos destes últimos são elementos do tipo “Identificador da Manifestação” e “Linguagem da Expressão”.

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

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

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

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

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

RDF apoiam a construção de vocabulário e dos relacionamentos para a comunidade da *Web Semântica*.

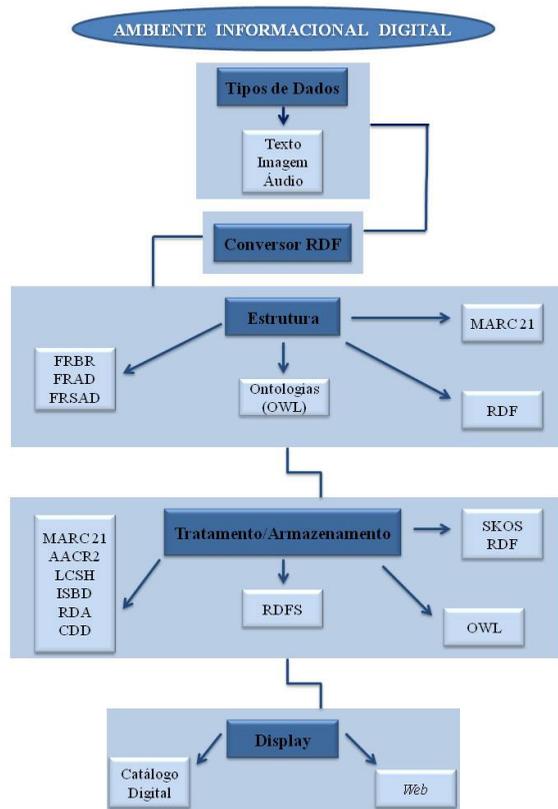
O desenvolvimento de RDA implica para as bibliotecas condições de avançar rapidamente para um serviço de intercâmbio e de interoperabilidade mais amplo, reutilizando os dados bibliográficos de outros ambientes informacionais digitais. Os elementos e vocabulários RDA fornecem a base para a migração do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

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

4 PROPOSTA DE REQUISITOS FUNCIONAIS PARA MODELAGEM ESTRUTURAL DE AMBIENTES INFORMACIONAIS DIGITAIS E INTEROPERÁVEIS

Como um ponto de partida para garantir a interoperabilidade, esta investigação, a partir das observações e reflexões no campo da Catalogação Descritiva, destaca alguns requisitos e diretrizes funcionais que podem ser utilizados no estabelecimento da interoperabilidade em ambientes informacionais digitais de uma forma mais efetiva. Para tanto, pretende-se compreender a estrutura intangível, numa proposta apresentada em camadas sobrepostas, conforme a figura 1.

FIGURA 1: Diretrizes funcionais para a interoperabilidade.



fonte: CASTRO (2012, p. 173).

Destaca-se que a proposta de requisitos funcionais para a modelagem estrutural deve trabalhar em sinergia para a consistência e o funcionamento pleno do ambiente informacional digital. Castro (2012) apresenta o detalhamento em cada camada com suas especificações, como segue:

- **Tipologia dos dados:** nessa fase inicial o projetista (catalogador) define quais os dados serão trabalhados para alimentação e modelagem do ambiente informacional, a partir do recurso bibliográfico a ser catalogado. Exemplo: dados textuais, imagéticos, áudio etc. Vale destacar que nessa tese nos teremos apenas nos dados textuais, explicitados nos códigos de catalogação (AACR2) e nos padrões de metadados (MARC 21).
- **Preparação dos dados:** uma vez definidos os dados bibliográficos a serem utilizados

no sistema, a preparação dos dados consiste na adoção de ferramentas para a conversão⁵ de dados em RDF. Os dados que foram extraídos de outras fontes, ou seja, dados não RDF, o W3C tem recomendado alguns conversores que auxiliam a transformação, como RDFizer. A adoção do RDFizer se justifica uma vez que não se tem oficialmente representações de dados MARC em RDF, nesse sentido, tal conversor transforma dados do padrão de metadados MODS em RDF, iniciativa essa promovida pela *Library of Congress* (MODS, 2011).

- **Tratamento e Armazenamento dos dados:** Após a conversão dos dados em RDF, a próxima camada consiste na adoção efetiva pelo catalogador das regras e/ou dos esquemas de descrição bibliográfica (AACR2 e RDA), ou seja, a catalogação dos recursos bibliográficos, na confecção padronizada de metadados; a definição dos padrões de metadados (MARC 21), da arquitetura de metadados RDF para a estruturação dos dados e RDF *Schema* para sua validação. Vale dizer que nesse momento o catalogador deverá também adotar os Requisitos Funcionais para Registros Bibliográficos (FRBR), em consonância com os Requisitos Funcionais para Dados de Autoridade (FRAD) e os Requisitos Funcionais para Dados de Assunto (FRSAD). As ontologias aparecem nesse contexto para definir os conceitos dos elementos de um registro bibliográfico, por meio das regras e esquemas de descrição para a confecção metodológica dos metadados e dos padrões de metadados.
- **Apresentação (Display):** A fase final consiste em tornar disponíveis os dados (*output*) e apresentá-los aos usuários do ambiente informacional. Os dados poderão aparecer da maneira como foram construídos e armazenados (*input*) nas camadas 1 e 3, para a camada tangível de recuperação e também visualizado na *Web*.

⁵ "Um conversor para RDF é uma ferramenta que converte aplicações de dados de um formato específico em RDF para uso com ferramentas de RDF e integração com outros dados. Conversores podem ser parte de um esforço de migração, ou parte de um sistema em execução que fornece uma visão Web Semântica de uma determinada aplicação". (W3C, 2014, tradução nossa).

Acredita-se que esses requisitos e recomendações podem propiciar uma modelagem dos catálogos melhor estruturados, para posterior recuperação, uso e re (uso) das informações, garantindo a interoperabilidade e potencializando os relacionamentos bibliográficos semânticos, iniciativa essa que vai de encontro aos ideais vislumbrados pela *Web Semântica*.

5 À GUIA DE REFLEXÕES

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.

Para Castro (2012), o formato de metadados 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.

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 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.

Acredita-se que tais requisitos funcionais propostos nesse trabalho, podem 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*.

INTEROPERABILITY'S ELEMENTS IN THE PERSPECTIVE DESCRIPTIVE CATALOGING

ABSTRACT

In the digital informational environment, currently, where various kinds of resources coexist with shapes and patterns of heterogeneous metadata, many efforts have been made to achieve interoperability in order to use multiple metadata standards and reuse the metadata records through the development strategies, ranging from simple mappings between metadata elements to modeling complex structural. The intensive use of Information and Communication Technologies (ICT) generates the development of a various possibility of metadata standards, essential in environments manipulation of bibliographic data. The objective is to point out to what extent the methods of Descriptive Cataloging, explicit codes of cataloging (AACR2 and RDA) and metadata standards, favor the design model of digital informational environments for better structured information retrieval. For this purpose, a methodology of analysis of the exploratory and descriptive literature with the study of scientific production of the main area of Information Science and Librarianship, it is possible to identify and understand the major technological changes in the heart of Descriptive Cataloging in the digital scope. As a result is presented a proposal for modeling a functional requirements of digital informational environments that ensure interoperability in the bibliographic domain. It is believed that the use of metadata, ontologies and the theoretical and methodological foundations of Descriptive Cataloging, once articulated, can redesign options to new informational environments, whether in the modeling of the bibliographic catalogs, or in the form of representation of information resources, ensuring interoperability.

Keywords: *Metadata standards. Descriptive cataloging. Interoperability. Structural Modeling of digital informational environments. Information and Technology.*

Artigo recebido em 08/11/2013 e aceito para publicação em 19/07/2014

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.
- ALVES, R. C. V.; SANTOS, P. L. V. A. C. **Metadados no domínio bibliográfico**. Rio de Janeiro: Intertexto, 2013. 196 p.
- ALVES, R. C. V. **Web Semântica: uma análise focada no uso de metadados**. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia Ciências, Universidade Estadual Paulista, Marília, 2005.
- BLATTMANN, U.; SILVA, F. C. C. da. Colaboração e interação na web 2.0 e biblioteca 2.0. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 12, n. 2, p. 191-215, jul./dez., 2007.
- CASTRO, F. 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.
- DELSEY, T. **RDA: resource description and access**. 2010. Disponível em: <<http://www.rda-jsc.org/rda.html>>. Acesso em: 25 de fev. 2014.
- GODBY, C. J.; SMITH, D.; CHILDRESS, E. Two paths to interoperable metadata. In: **Proceedings of DC-2003: Supporting Communities of Discourse and Practice**. Metadata Research & Application. Seattle, Sep./Oct. 2003. Available: <http://www.siderean.com/dc2003/103_paper-22.pdf>. Access to: 20 jan. 2014.
- GUY, M. **Interoperability focus: looking at interoperability**. Ukoln, 2005. Available: <<http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>>. Access to: 20 mar. 2014.
- HILLMANN, D. et al. RDA vocabularies: process, outcome, use. **The Magazine of Digital Library Research**, Vol. 16, no. 1/2, Jan. /Feb. 2010.
- IEEE. STANDARDS INFORMATION NETWORK. **The authoritative dictionary of IEEE standards terms**. 7 ed. New York: IEEE, 2000.
- IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional requirements for bibliographic records: final report. **UBCIM Publications - New Series**, vol. 19. München: K. G. Saur, 1998. 136 p. Available: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>. Access to: 18 dez. 2013.
- LYTRAS, M. D; SICILIA, M. A. Where is the value of metadata? **International Journal of Metadata, Semantics and Ontologies**, Vol. 2, no. 4, p. 235-241, 2007.
- MARINO, M. T. **Integração de informações em ambientes científicos na web: uma abordagem baseada na arquitetura RDF**. 2001. Dissertação (Mestrado em Informática) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.
- RODRÍGUEZ, E. M. **Metadados y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales**. Gijón: Trea, 2002. 429 p.
- MODS. **Metadata Object Description Schema**. Library of Congress, Mar. 2011. Available: <<http://www.loc.gov/standards/mads/rdf/>>. Access to: 21 maio 2014.
- NILSSON, M. **From interoperability to harmonization in metadata standardization: designing an evolvable framework for metadata harmonization**. 2010. 125 f. Thesis (Doctoral in Computer Science and Communication) - KTH School of Computer Science and Communication, Stockholm, 2010.
- SANTOS, P. L. V. A. C.; CORRÊA, R. M. R. **Catalogação: trajetória para um código internacional**. Niterói: Intertexto, 2009. 80 p.

SANTOS, P.; L. V. A. C.; ALVES, R. C. V. Metadados e web semântica para a estruturação da web 2.0 e web 3.0. **Datagramazero - Revista de Ciência da Informação**, v. 10, n. 6, dez. 2009.

SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **TransInformação**, Campinas, v. 20, n. 2, p.133-148, maio/ago. 2008.

TAYLOR, A. **Understanding FRBR**: what it is and how it will affect our retrieval tools. Westport: Libraries Unlimited, 2007. 186 p.

WORLD WIDE WEB CONSORTIUM (W3C). **OWL 2 Web Ontology Language, Structural Specification and Functional-Style Syntax**, W3C Recommendation, Oct. 2009. Available: <<http://www.w3.org/TR/owl-syntax/>>. Access to: 20 jan. 2014.

W3C. **Converter to RDF**. 2014. Available: <<http://www.w3.org/wiki/ConverterToRdf>>. Access to: 21 jan. 2014.

Agradecemos à *Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)*, pelo apoio financeiro fornecido ao desenvolvimento dessa pesquisa.