

# Tecnologia, Metadados e Interoperabilidade: aspectos na modelagem de ambientes informacionais digitais

FABIANO FERREIRA DE CASTRO

*Programa de Pós-Graduação em Ciência da Informação  
(UFSCar), Brasil.*

## INTRODUÇÃO

A preocupação da comunidade biblioteconômica e da Ciência da Informação advinda com as novas regras para a descrição bibliográfica aplicadas ao 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.

Nas 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 tecnologias atuais disponíveis 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 problemas como a padronização das formas de representação da informação 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.

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 pontos de reflexão, destaca-se em que medida os atuais ambientes informacionais digitais precisam ser repensados, a partir da modelagem conceitual e estrutural dos dados, na camada intangível, definida na representação e na descrição dos dados bibliográficos, com alcance à interoperabilidade.

## APONTAMENTOS E IMPACTOS TECNOLÓGICOS NO CENÁRIO DA CATALOGAÇÃO DESCRITIVA

Percebe-se que 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. Dessa forma, questiona-se se o novo formato para a descrição e a representação dos dados será XML (eXtensible Markup Language), ele fará uso de

RDF (Resource Description Framework) e padrões de dados ligados<sup>1</sup>

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

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

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

---

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 e Bizer, 2011, tradução nossa).

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, *ibid.*).

A ligação pode também ser alcançada através dos relacionamentos verticais, similares aos termos (mais amplos e mais restritos) num tesauro. 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 (*ibid.*) afirma que infelizmente isso não é possível devido à forma que ISBD e RDA foram definidos em RDF.

Coyle (*ibid.*) 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 (*ibid.*) 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 (*ibid.*) 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, *ibid.*).

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 dos esquemas de descrição oriundos da catalogação, porém não explicitados aos sujeitos institucionais para o uso e (re) uso das informações nos ambientes de manipulação de dados bibliográficos e catalográficos.

## DA HETEROGENEIDADE DOS DADOS AOS METADADOS: CONSTRUCTOS ACERCA DA INTEROPERABILIDADE

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

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

Um objetivo comum desses esforços é permitir que diferentes comunidades, com diferentes tipos de informação e usando diferentes tecnologias, consigam um nível geral de compartilhamento de informação e, por meio de processos de agregação apoiados por tecnologia da informação, criem novos e mais poderosos tipos de serviços de informação. (Sayão e Marcondes, 2008). De acordo com Méndez Rodríguez (2002, p. 312, tradução nossa),

A interoperabilidade supõe complementar uma série de estratégias humanas ligadas à descrição consistente e flexível da informação mediante metadados, com um conjunto de ferramentas, *softwares* e protocolos que permitam os distintos sistemas de informação

(bibliotecas digitais) estabelecer mecanismos de conversão automática entre eles, e ainda, reconciliar suas diferenças e alcançar, ao menos, um acesso universal integrado aos recursos informacionais.

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

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

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.

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

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

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

Como podemos definir conceitualmente interoperabilidade no contexto dos metadados? O IEEE (2000) diz que a interoperabilidade corresponde à habilidade de dois ou mais sistemas ou componentes trocar informações e utilizá-las neste processo. Nilsson (2010) diz que a interoperabilidade de metadados consiste na capacidade de dois ou mais sistemas, ambientes informacionais ou componentes trocarem dados descritivos sobre recursos 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, p. 136),

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

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 e Sicilia, 2007). Por essa razão, a administração e o intercâmbio de metadados é a atividade central em muitos sistemas que gerenciam os objetos digitais e não digitais, tais como os sistemas de gerenciamento de conteúdo, repositórios de objetos de aprendizagem e bibliotecas (digitais ou convencionais).

Nilsson (2010) acrescenta que especificações e padrões de metadados acrescentam valores adicionais ao reduzir o limiar para o desenvolvimento de sistemas que realizam intercâmbio, reusam e combinam metadados de diferentes fontes. Um padrão comum garante uma melhor documentação e um melhor acesso às ferramentas reutilizáveis. Este é o valor central para a interoperabilidade de metadados. Percebendo o potencial inerente ao uso de metadados interoperáveis informado requer uma coordenação em grande

escala entre os atores relevantes em um campo de prática e aplicação. Especificações de metadados tendem a ser concebidos para uma determinada comunidade particular, com mais ou menos itens bem definidos e descritos e cenários comuns de uso.

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

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

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

Godby, Smith e Childress (2003) argumentam baseados em experimentos com o cruzamento de metadados (*crosswalks*), que as traduções completas de um padrão para outro são possíveis numa comunidade de prática específica, enquanto somente traduções

parciais são possíveis entre elas. Os autores exemplificam os metadados de uma biblioteca que podem ser combinados com outras especificações de bibliotecas diferentes, desde que usando um padrão comum para a descrição dos recursos bibliográficos.

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

Subjacentes aos formatos e 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.

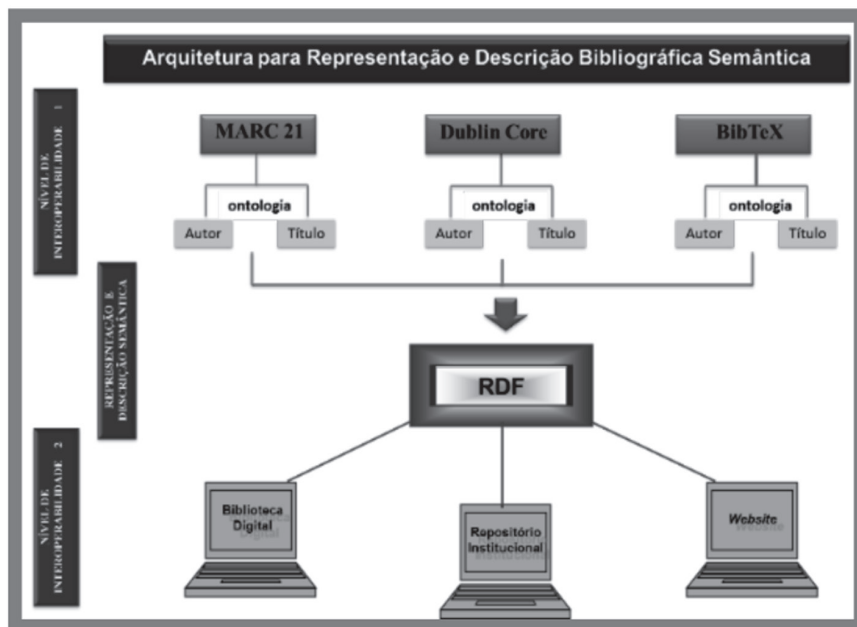
## PROPOSTA DE MODELAGEM ESTRUTURAL DE DADOS NOS DOMÍNIOS BIBLIOGRÁFICO E WEB

De acordo com os apontamentos e reflexões acerca das tecnologias digitais, e com respaldo na literatura científica, verifica-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.

Castro (2012), 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, 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*.

*Figura 1.* Modelo para a Representação e a Descrição Bibliográfica semântica e níveis de interoperabilidade.



Fonte: Castro (2012, p. 178).

A arquitetura para a Representação e a Descrição Bibliográfica Semântica vislumbra a possibilidade da construção de representação e de descrição bibliográfica semântica e a interoperabilidade semântica entre, não somente padrões de metadados, mas também entre ambientes e sistemas informacionais digitais heterogêneos.

Na parte superior da arquitetura, têm-se alguns padrões de metadados (MARC21, Dublin Core e BibTeX), descrevendo alguns

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

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

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).

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

Vale destacar nesse momento que a arquitetura supracitada apresenta ontologias para apenas dois elementos da descrição bibliográfica na estrutura MARC 21 (autor e título), uma vez que é necessário o desenvolvimento e a criação de ontologias para determinar a semântica de todos os metadados do padrão MARC 21, pois quanto mais complexa e completa for a estruturação, haverá uma compreensão mais intuitiva e efetiva pela máquina promovendo a interoperabilidade.

Ainda na fase de confecção dos metadados, a atribuição das ontologias 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.

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, Website etc., definidos no modelo, sendo que a representação e a descrição dos recursos informacionais, podem estar ou não contemplados num mesmo padrão. O nível de interoperabilidade 2 objetiva a possibilidade de interoperabilidade semântica entre múltiplos ambientes e sistemas informacionais digitais e o armazenamento das descrições bibliográficas semânticas numa única plataforma/interface.

Ressalta-se, que quanto mais completa e complexa for uma descrição, melhores possibilidades nos níveis de interoperabilidade semântica haverá, com a possibilidade de uso, de preservação e de (re)

uso das informações de forma mais efetiva pelo usuário final, multidimensionando as formas de acesso aos recursos informacionais.

Portanto, acredita-se que tal modelo 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*.

## À GUIA DE CONSIDERAÇÕES E REFLEXÕES

A utilização de ontologias, de metadados e dos fundamentos teóricos e metodológicos da Catalogação Descritiva uma vez articuladas, podem redesenhar novas opções aos ambientes informacionais, seja na modelagem do banco de dados (catálogos), ou ainda, na forma de representação dos recursos informacionais, garantindo a possibilidade de interoperabilidade semântica e no auxílio às formas de apresentação dos recursos, multidimensionando as formas de busca aos usuários, de acesso e de recuperação de informações pertinentes, significativas e relevantes, bem como o uso, a preservação e o (re) uso numa única interface. (Castro, 2016, tradução nossa).

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.

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.

O modelo para a representação e a descrição bibliográfica semântica e os níveis de interoperabilidade desenvolvido nesse estudo, 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 BIBLIOGRÁFICAS

- AACR2. (2005). *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). São Paulo: FEBAB, 2005.
- Alves, R. C. V. (2005). 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.
- Castro, F. F. de. (2016). Functional requirements for Bibliographic Description in Digital Environments. *TransInformação*, 28: 223-231.
- . (2012). Elementos de interoperabilidade na catalogação descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais. Tese (Doutorado em Ciência da Informação). Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.
- Coyle, K. (2012). *Bibliographic Framework: RDF and linked data*. Berkeley, United States. Disponível em: <http://kcoyle.blogspot.com/2012/01/bibliographic-framework-rdf-and-linked.html>. Acesso em: 20 jan. 2016.
- . (2011). MARC 21 as Data: a Start. *Code4Lib Journal*, 14.
- Godby, C. J., D. Smith e E. Childress. (2003). Two Paths to Interoperable Metadata. *Proceedings of DC-2003: Supporting Communities of Discourse and Practice. Metadata Research & Application*. Seattle, Sep./Oct. Disponível em: [http://www.siderean.com/dc2003/103\\_paper-22.pdf](http://www.siderean.com/dc2003/103_paper-22.pdf). Acesso em: 20 jan. 2016.

- Guy, M. (2005). Interoperability Focus: Looking at Interoperability. Ukoln. Disponível em: <http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>. Acesso em: 20 mar. 2016.
- Heath, T. e C. Bizer. (2011). *Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool, 1* (1). (Synthesis Lectures on the Semantic Web: Theory and Technology). Disponível em: < <http://linkeddatabook.com/editions/1.0/#linkedData>>. Acesso em: 15 set. 2016.
- IEEE. (2000). Standards Information Network. *The Authoritative Dictionary of IEEE Standards Terms*. New York: IEEE.
- IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998) *Functional Requirements for Bibliographic Records*: final report. UBCIM Publications - New Series, vol. 19. München: K. G. Saur. Disponível em: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>. Acesso em: 18 jan. 2016.
- Lytras, M. D. e M.A. Sicilia. (2007). Where is the Value of Metadata? *International Journal of Metadata, Semantics and Ontologies*, 2 (4): 235-241.
- Méndez Rodríguez, E. (2002). *Metadados y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002.
- Nilsson, M. (2010). From Interoperability to Harmonization in Metadata Dstandardization: Designing an Evolvable Framework for Metadata Harmonization. Thesis (Doctoral in Computer Science and Communication). KTH School of Computer Science and Communication, Stockholm.
- Riley, J. (2010). *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 out. 2016.

- Santos, P. L. e M. R. Corrêa. (2009). *Catálogo: trajetória para um código internacional*. Niterói: Intertexto.
- Sayão, L. F. e C. H. Marcondes. (2008). O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, 20, (2): 133-148.
- Thomale, J. (2010). Interpreting MARC: Where's the bibliographic data? *Code4Lib Journal*, 11, 2010.
- World Wide Web Consortium (W3C). (2009). *OWL 2 Web Ontology Language, Structural Specification and Functional-Style Syntax*, W3C Recommendation. Disponível em: <<http://www.w3.org/TR/owl-syntax/>>. Acesso em: 20 jan. 2016.