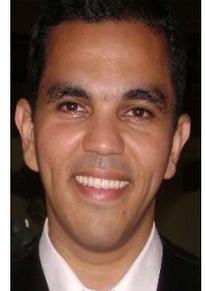


Modelos conceituais e representação descritiva da informação



Elisabete Gonçalves de Souza
Universidade Federal Fluminense
elisabetes.souza@gmail.com



Wellington Freire Cunha Costa
Universidade Federal Fluminense
freirew@yahoo.com.br

1 Introdução

Os metadados são essenciais para a descrição documental pelo fato de transformarem a informação em códigos auxiliando na sua representação, recuperação e acesso. Os repositórios digitais fazem uso dos metadados para descrever e tornar possível a busca e a recuperação da informação em linha, possibilitando a interoperabilidade das bases de dados.

De acordo com Marcondes (2005, p.96):

Um dos maiores objetivos do uso de metadados no contexto da Web é permitir não só descrever documentos eletrônicos e informações em geral, possibilitando sua avaliação de relevância por usuários humanos, mas também permitir agenciar computadores e programas especiais, robôs e agentes de *software*, para que eles compreendam os metadados associados a documentos e possam então recuperá-los, avaliar sua relevância e manipulá-los com mais eficiência.

Para garantir a integridade das descrições e interoperabilidade dos sistemas é primordial o uso de normas e padrões. Para os usuários essa ação permite-lhes localizar e escolher dentre as várias expressões e manifestações de um recurso, aquela que melhor atenda às suas necessidades de informação. Para os gestores, significa compartilhar os registros, usar e reusar as informações já processadas contextualizando-as de acordo com a política de seu sistema de informação.

O Dublin Core (DC) é o padrão para a estruturação de metadados mais utilizado para a criação de repositórios e apresenta as seguintes características: a) flexibilidade –

os elementos que o compõem são opcionais, podendo ser exibidos aleatoriamente e se houver necessidade podem ser repetidos ou modificados por qualificadores; b) simplicidade – é de fácil manuseio, sendo autoexplicativo, permitindo que o próprio autor do documento faça a descrição sem a necessidade de intensos treinamentos; c) extensibilidade – por ser um modelo simples e flexível, permite que diversas comunidades em áreas distintas utilizem o padrão DC trocando informações e obtendo o acesso a elas; d) interoperabilidade – um modelo comum de descrição utilizado por diversas áreas proporciona aos usuários maior habilidade nas pesquisas; e) âmbito internacional – embora tenha sido desenvolvido em inglês, já existem versões em outras línguas; infraestrutura adequada e escopo internacional (DCMI, 2015).

Conforme Marcondes (2005) o padrão DC foi pensado para ser simples o suficiente e autoexplicativo de modo que o próprio autor do documento possa descrevê-lo ao publicá-lo eletronicamente, o que seria impossível em outro padrão como o MARC, por exemplo, pois exige o aporte de profissionais de informação treinados dada a sua complexidade. Indo em direção oposta,

o Dublin Core, pode ser definido como sendo o conjunto de elementos de metadados planejado para facilitar a descrição de recursos eletrônicos. A expectativa é de que os autores e Websites, que não possuam conhecimentos em catalogação, possuam capacidade de usar o Dublin Core para descrição de recursos eletrônicos, tornando suas produções mais visíveis aos mecanismos de busca e sistemas de recuperação (SOUZA; VENDRUSCULO; MELO, 2000, p. 93).

O DC é um padrão que complementa os demais, pela sua simplicidade na descrição dos registros, funcionando como um “meio-termo” entre dois tipos de descrição: as automáticas realizadas pelos Serviços de Indexação e Resumo e as catalográficas produzidas pela biblioteca. Por ser baseado no protocolo Z39.50 (hoje reestruturado como SRU - *Search and Retrieve URL*, e SRW - *Search/Retrieve Web Service*), permite a troca de informações entre diferentes sistemas de bases de dados, busca simultânea e importação de registros, podendo ser implantado em qualquer plataforma.

A interoperabilidade das descrições feitas no DC está garantida pelo uso da linguagem XML (*Extensible Markup Language*) e da semântica da RDF (*Resource*

Description Framework). De acordo com Marcondes (2005, p. 108) “enquanto a XML é uma linguagem genérica para a descrição da estruturação de documentos eletrônicos, RDF é específica para criar metadados com a finalidade de localizar e identificar recursos. A codificação em RDF permite descrever e pesquisar recursos de forma mais flexível, mediante representação de relacionamentos entre estes, na forma de triplas compostas por objeto-atributo-valor, sendo que um valor pode ser um literal ou outro recurso” (CAMPOS; CAMPOS; CAMPOS, 2005, p. 66). Ou seja: por meio da interoperabilidade semântica, o formato DC pode ser utilizado por diversas áreas e por ser flexível permite a extensibilidade e modularidade dos metadados, tornando possível a troca de informações entre as comunidades científicas de forma fácil e rápida.

2 Modelo entidade-relacionamento

O novo modelo conceitual FRBR (*Requisitos Funcionais para Registros Bibliográficos*) foi inspirado no metamodelo Entidade-Relacionamento (E-R) criado na década de 1970 pelo cientista da computação Peter Chen para otimizar as rotinas dos bancos de dados de modo que os registros gerados nos bancos fossem interligados, apurando a recuperação de itens importantes de forma fácil e rápida. Segundo Moreno (2006, p.34):

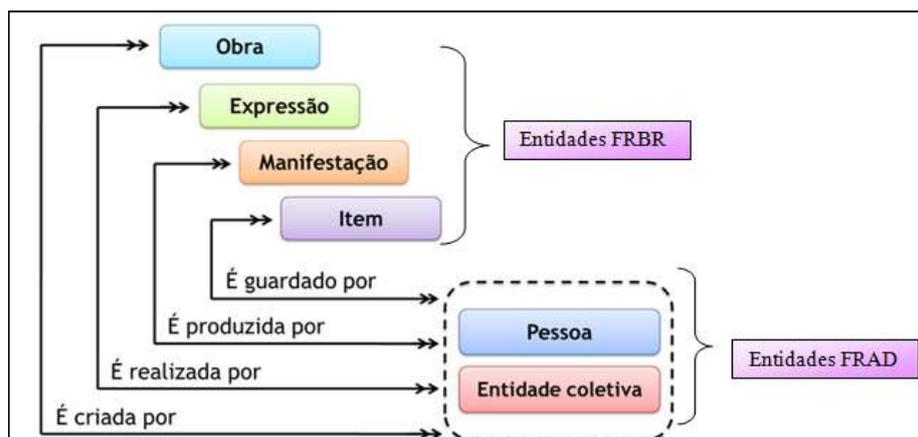
Modelos de dados são um conjunto de conceitos utilizados para descrever um banco de dados. O modelo E-R é um modelo lógico com base em objetos, e a identificação de entidades e relacionamentos é entendida como a captura da semântica dos dados, para projetar um banco de dados. Este modelo não visa à implementação e sim à modelagem/representação dos dados. A partir desta modelagem, é possível implementar um banco de dados em outros modelos de dados: orientado a objeto, relacional, etc.

Segundo Silva e Santos (2012) a modelagem conceitual descreve que um domínio é composto por coisas representadas pelas entidades, que por sua vez possuem características a serem descritas, denominadas atributos; e pelos relacionamentos existentes entre elas. Por isso as bases de dados relacionais possuem maior capacidade semântica.

A figura abaixo ilustra o modelo E-R em um banco de dados com informações sobre um conjunto de objetos classificados como pessoa (entidade FRAD), um conjunto

de objetos classificados como manifestação (entidade FRBR), um conjunto de associações, cada uma ligando uma manifestação a uma pessoa – criada por (relacionamento) e um conjunto de informações que caracterizam as entidades (atributos).

Figura 1 – Modelo E-R (domínio bibliográfico)



Fonte: Assumpção (2012). Adaptação dos autores

O modelo (E-R) foi criado para melhorar o desenvolvimento dos bancos de dados fornecendo buscas mais precisas para os usuários. O modelo conceitual FRBR foi criado pela IFLA (*International Federation of Library Associations and Institutions*) também com esse propósito: representar as necessidades do usuário de catálogos e bibliografias levando em consideração suas tarefas ou funções - encontrar, identificar, selecionar e obter - com o objetivo de aproximá-lo da obra.

O modelo FRAD (*Functional Requirements for Authority Data*) é uma extensão dos FRBR para registro de autoridade, corresponde ao cabeçalho autorizado para uma entidade estabelecido de acordo com as regras de uma agência catalogadora. No universo bibliográfico uma entidade FRAD arrola informações sobre uma pessoa ou entidade, a qual tem seu nome usado como ponto de acesso controlado nos registros bibliográficos, seja em um catálogo ou em um repositório.

As entidades FRAD são responsáveis pela criação, produção, edição e outras responsabilidades relacionadas a uma obra. Seu uso amplia não só a capacidade de relacionamentos como a de associações de dados, já que ao individualizarmos uma entidade pela aplicação de atributos, podemos associá-la a dados abertos, como

currículo lattes, grupos de pesquisa, etc. principalmente para autores acadêmicos com produção em repositórios digitais.

2.1 Aplicação do modelo FRAD em repositório

O uso do padrão DC associado ao modelo E-R possibilitará que o usuário/autor ao depositar um documento possa descrevê-lo de forma mais precisa, já que o novo padrão RDA (*Resource Description and Access*) permite a descrição dos metadados bibliográficos e de autoridade de modo extensivo, em linguagem próxima daquela usada pelos usuários ao definir as entidades e seus atributos.

Para exemplificar usamos como campo empírico o repositório institucional LUME da Universidade Federal do Rio Grande do Sul. Para iniciarmos o processo de modelagem desse repositório contextualizamos sua estrutura aproximando-a das entidades da família FRBR. Para efeito de estudo, faremos a modelagem das entidades FRAD, acrescentando atributos que identificam e qualificam as pessoas responsáveis pela produção da obra (uma dissertação de mestrado) e a entidade/instituição de pesquisa com a qual se relaciona.

Figura 2 - Identificação das entidades

Atributo/Entidade FRBR	Título	"Gestor de quem?" A contribuição do bibliotecário para a gestão de pessoas da organização : um estudo de caso
Entidades - FRAD	Autor	Soares, Monica Fonseca
Atributo/Entidades FRBR	Orientador	Nakayama, Marina Keiko
Entidade - FRAD	Data	2007
Entidades FRAD	Nível	Mestrado
	Instituição	Universidade Federal do Rio Grande do Sul. Escola de Administração. Programa de Pós-Graduação em Administração.
	Assunto	Bibliotecário : Perfil Gerente Gestão de pessoas Recursos humanos [en] Librarian [en] Library management [en] Manager [en] Personnel management

Fonte: elaboração dos autores

A simulação abaixo demonstra a modelagem da estrutura descritiva do Repositório Institucional (RI) LUME, campos referentes aos dados de autoridade. A adição de atributos FRAD aos elementos DC contributor.advisor e contributor.author que

representam pontos de acesso/pessoa da obra "*Gestor de quem?*" *A contribuição do bibliotecário para a gestão de pessoas da organização : um estudo de caso*" visa aprimorar processo de busca e recuperação da informação no sentido que permite aos usuários do repositório novas possibilidades de relacionamentos durante a pesquisa. Com isso, o usuário terá mais informações sobre o criador do documento/recurso e os diferentes relacionamentos que este mantém com outras entidades, no caso de nosso exemplo, seu orientador, o programa no qual está filiado, grupos de pesquisa, etc.

Nessa direção, podemos dizer que o acréscimo de novos atributos (currículo lattes, filiações institucionais, etc.) aos metadados de autoridade identificados no registro DC (autor, orientador) amplia os processos de busca e recuperação de informação; levam o pesquisador a navegar para além do conteúdo da obra à medida que a contextualizava em relação aos seus pares. Para recursos como as dissertações e teses, tais relacionamentos são fundamentais para a interlocução entre pesquisadores o que faz dos repositórios uma poderosa ferramenta de comunicação científica.

No caso do repositório LUME, para que as relações fossem mais completas seria interessante a descrição dos metadados dos membros da banca. No entanto, essa informação foi suprimida, porque o gestor, a Universidade Federal do Rio Grande do Sul, segue para esse tipo de obra (teses e dissertações) os metadados estruturais (MTD-BR) definidos pela BDTD (Biblioteca Digital Brasileira de Teses e Dissertações), pois integram esse consórcio coordenado pelo IBICT.

A figura abaixo mostra a aplicação do modelo conceitual FRAD na estrutura descritiva do RI-LUME referente aos dados de autoridade (entidade pessoa) com o acréscimo de atributos que identificam as autoridades autor e orientador de forma contextualizada, no que diz respeito à sua inserção acadêmica. Além da forma extensível, adotamos também cabeçalhos para as entidades relacionadas aos nomes, por ser esta a forma normalizada predominante na maioria dos catálogos, o que nos permite interligar os dados sobre as diferentes comunidades científicas.

Figura 3 – Aplicação do FRAD no RI-LUME

Entidade FRAD	dc.contributor.advisor	Nakayama, Marina Keiko
Atributo FRAD	Link Currículo Lattes	http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4798728D2
Atributo FRAD	Linha de pesquisa	Gestão do conhecimento e compartilhamento
Atributo FRAD	Grupo de pesquisa	Núcleo Estudos e Observação de Gestão, Aprendizagem e Pessoas
Atributo FRAD	Link grupo/pesquisa	http://www.neogap.ufsc.br/wp/?page_id=58
Entidade FRAD	dc.contributor.author	Soares, Monica Fonseca
Atributo FRAD	Link Currículo Lattes	http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4782635D0
Atributo FRAD	Programa/Pós-graduação	Programa de Pós-Graduação em Administração
Atributo FRAD	Forma normalizada do nome do Programa	Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Administração
Atributo FRAD	Link para o programa	http://www.ufrgs.br/escoladeadministracao/pos-graduacao/

Fonte: elaboração dos autores

Na institucionalização de um campo científico é necessário que se tenha confiabilidade no conhecimento produzido. Um repositório institucional é um meio de divulgação do conhecimento e resultados advindos das pesquisas empreendidas pelos cientistas; um canal de comunicação que arrola diferentes produções formais e informais cujos resultados obtidos podem ser acessados e compartilhados por outros pesquisadores. Por isso, para além da flexibilidade proporcionada pela aplicação do modelo FRBR, cabe ao gestor acuidade quanto à descrição e codificação dos metadados, levando em consideração o uso da RDA e da linguagem RDF, aferindo assim, as vantagens advindas do uso desses padrões na recuperação e acesso a informações em ambiente digitais.

3 Considerações parciais

Como vimos, um repositório institucional tem o objetivo de preservar o conhecimento acumulado nas instituições apresentando como principais vantagens o armazenamento de grande quantidade de informações, facilidade de busca e recuperação e autonomia para o produtor fazer a descrição e o autoarquivamento de seus documentos.

O processo de simulação da aplicação dos atributos do FRAD na estruturação descritiva, demonstrou que tal implementação melhorará consideravelmente o processo de busca e recuperação da informação, qualificando os registros com o acréscimo de atributos individuais como o currículo lattes e de contextualização acadêmica, como a vinculação aos grupos e/ou núcleos de pesquisa, o que permitirá ao usuário uma visão

mais abrangente sobre as comunidades científicas nas quais se inserem os autores e seus pares.

No que diz respeito à recuperação da informação dois aspectos merecem destaques: a) o detalhamento de um registro com acréscimos de atributos amplia as possibilidades de relacionamentos entre as entidades pessoas e as demais entidades bibliográficas que compõem o registro DC; b) a incorporação de links (URL) como o currículo lattes inaugura a possibilidade de associar as ferramentas bibliográficas aos dados abertos disponíveis na web.

Referências

- ASSUMPCÃO, Fabrício. **O que é FRBR?** Disponível em: <<http://fabricioassumpcao.com/2012/07/o-que-e-frbr.html>>. Acesso em: 4 maio 2015.
- CAMPOS, M. Luiza Machado; CAMPOS, M. Luiza de Almeida; CAMPOS, Linair Maria. Web semântica e a gestão de conteúdos informacionais. In: MARCONDES, Carlos Henrique et al. (Org.). **Bibliotecas Digitais: saberes e práticas**. Salvador, EDUFBA; Brasília; IBICT, 2005. p. 55-74.
- DCMI. **Using Dublin Core**. Disponível em: <http://dublincore.org/documents/usageguide/> Acesso 30 abr. 2015.
- MARCONDES, Carlos Henrique. Metadados: descrição e recuperação na Web In: ____ et al. (Org.). **Bibliotecas digitais: saberes e práticas**. Salvador, BA: EDUFBA; Brasília; IBICT, 2005. p. 77-143
- MORENO, Fernanda Passini. **Requisitos funcionais para registros bibliográficos – FRBR: um estudo no catálogo da Rede Bibliodata**. 2006. 199 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, 2006.
- SILVA, Renata Eleutério da; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Requisitos Funcionais para Registros Bibliográficos (FRBR): considerações sobre o modelo e sua implementabilidade. **RBD**. São Paulo, v. 8, n. 2, p.116-129, Jul./Dez. 2012.
- SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, v. 29, n. 1, p. 93-102, Jan./Abr. 2000.
- UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL. **Repositório Lume**. 2015. Disponível em: <http://www.lume.ufrgs.br> Acesso 30 abr. 2015.