

BIBFRAME: tendência para a representação bibliográfica na Web

Felipe Augusto Arakaki (IFSP/UNESP) - arakaki@reitoria.unesp.br

Luiz Felipe Galeffi (UNESP) - luizgaleffi@gmail.com

Rachel Cristina Vesu Alves (FFC - UNESP) - alves.rachelcv@gmail.com

Plácida Amorim da Costa Santos (UNESP) - placidasantos@gmail.com

Resumo:

O uso de tecnologias sempre esteve presente no desenvolvimento da Biblioteconomia, particularmente no âmbito da descrição de recursos. Esta questão pode ser observada logo no início da década de 1960, com o desenvolvimento do Machine-Readable Cataloging (MARC). Entretanto, o formato MARC 21 permaneceu sem amplas modificações em sua estrutura para o compartilhamento de dados, excetuando-se a disponibilização do MARCXML em 2002, que permitiu apenas uma nova possibilidade de favorecer o intercâmbio de dados com outros formatos. Neste contexto foi desenvolvido pela LC, em maio de 2011, o Bibliographic Framework (BIBFRAME), aproveitando-se de novas tecnologias como o Resource Description Framework (RDF) e os modelos conceituais Functional Requirements for Bibliographic Records (FRBR) e Functional Requirements for Authority Data (FRAD), bem como da ideia de inserção de dados bibliográficos na web a partir dos princípios do Linked Data. Dessa forma, o objetivo desse trabalho é apresentar o BIBFRAME como novo instrumento para catalogação. É caracterizada por uma pesquisa qualitativa e exploratória. Nesse contexto, é apresentada uma breve contextualização da evolução do MARC, MARCXML e posteriormente, o BIBFRAME. Conclui-se que a evolução dos padrões de metadados do domínio bibliográfico acarretará numa nova transição para os catálogos das bibliotecas. Baseados principalmente pelos modelos conceituais FRBR e FRAD, as diretrizes RDA e os conceitos da Web Semântica, tanto as estruturas dos catálogos, quanto o processo de catalogação se modificarão. Essas mudanças são fundamentais para maximizar a busca e recuperação da informação dos usuários nos catálogos.

Palavras-chave: *BIBFRAME, Catalogação, Metadados, Linked Data*

Eixo temático: *Eixo 6: IV EEPC - Encontro de Estudos e Pesquisas em Catalogação. Organização e Tratamento da Informação: tecnologias e novas ferramentas, instrumentos, processos, produtos e serviços, políticas, cooperação.*

XXVII Congresso Brasileiro de Biblioteconomia e Documentação

Fortaleza, 16 a 20 de outubro de 2017.

1 INTRODUÇÃO

Devido à grande produção de informações registradas, em diversas áreas do conhecimento, a partir de meados do século XX, cresceu também a preocupação dos profissionais da área de Ciência da Informação em controlar esta produção e descrever todo esse volume de informação gerada. No contexto do domínio bibliográfico, foram criados os padrões de metadados baseados nos princípios e códigos de catalogação, como o formato *Machine-Readable Cataloging* (MARC) na década de 1960. Entretanto, com o desenvolvimento das novas tecnologias são exigidas atualizações para adequação à nova estrutura proposta para catalogação como as diretrizes - *Resource Description and Access* (RDA) - desenvolvido para substituir o *Anglo-American Cataloguing Rules* (AACR) e baseado nos modelos conceituais *Functional Requirements for Bibliographic Records* (FRBR) e *Functional Requirements for Authority Data* (FRAD). Sendo assim, a *Library of Congress* dos Estados Unidos (LC) lançou no ano de 2011, a *Bibliographic Framework Initiative*, com o intuito de apresentar uma estrutura que atenda às perspectivas da área de catalogação, chamado *Bibliographic Framework* (BIBFRAME), baseando-se nas tecnologias emergentes e das últimas décadas como o *Linked Data*.

Atualmente, o *World Wide Web Consortium* (W3C), consórcio internacional responsável pelo desenvolvimento de padrões para a *web*, tem desenvolvido iniciativas para o uso de padrões cada vez mais interligados e que possam trabalhar de forma a ampliar a interoperabilidade entre sistemas, disponibilizando assim as informações de um modo que os dados estejam ligados (*Linked Data*). Neste âmbito, tecnologias como a arquitetura de dados *Resource Description Framework* (RDF), os modelos conceituais FRBR e FRAD - desenvolvidos pela *International Federation of Library Associations* (IFLA) -, as diretrizes do RDA e a linguagem de marcação *eXtensible Markup Language* (XML), entre outros, fazem parte do desenvolvimento do BIBFRAME, possibilitando ao modelo uma adequação às perspectivas e tendências do tratamento da informação atual.

De acordo com Catarino e Souza (2012, p. 79), “[...] é importante ter uma grande quantidade de dados disponível em um formato padrão, acessível e gerenciável por tecnologias apropriadas.” Esta citação refere-se ao *Linked Data* e, mais precisamente, às bases da *Web Semântica*. Pensando também nas contribuições já trazidas pelas novas tecnologias, como o BIBFRAME, pela *Library of Congress*, tem como proposta ser “[...] o futuro da descrição bibliográfica que acontece dentro, entre e como parte da *web* e da rede mundial em que vivemos.” (LIBRARY OF CONGRESS, 2012, p. 3, tradução nossa).

Nesse contexto, o objetivo desse trabalho é apresentar o BIBFRAME como novo instrumento para catalogação de recursos informacionais, no contexto do ambiente da *Web*.

É caracterizada por uma pesquisa qualitativa e exploratória, pois, busca familiarizar e analisar as características do BIBFRAME e seu desenvolvimento. Para localização desses documentos (dissertações, teses, artigos, trabalhos de eventos, relatórios científicos e livros) foram utilizadas as bases de dados: Base de Dados Referencial de Artigos de Periódicos em

Ciência da Informação (Brapci), P@rthenon, Portal de Periódicos da Capes, *Scientific Electronic Library Online* (SciELO), na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Foram selecionados estudos que abordam a evolução do MARC e do desenvolvimento do BIBFRAME, além de textos considerados relevantes para explicitação deste trabalho.

2 Metadados bibliográficos

No domínio bibliográfico, os metadados tem uma função representativa mais específica do que no domínio *web*, e foram “[...] criados, aperfeiçoados e padronizados de acordo com os princípios, os códigos e as regras de catalogação [...]”. (ALVES; SANTOS, 2013, p. 85).

Segundo Assumpção e Santos (2015), o formato *Machine-Readable Cataloging* (MARC) surgiu na década de 1960, desenvolvido pela *Library of Congress* (LC), nos Estados Unidos, com intuito de criar registros catalográficos legíveis por máquina. Em 1966 foi criado o *MARC Pilot Project*, que em dois anos distribuiu cerca de cinquenta mil registros em fita magnética. Com a experiência bem-sucedida do projeto, em 1969 foi desenvolvido o MARC II, que deu origem a diferentes versões até chegar em sua última denominada MARC 21.

As diretrizes e padrões que estruturaram o formato MARC 21 são apenas adequações da prática profissional antes realizada pelos bibliotecários, baseada nos códigos de catalogação (*International Standard Bibliographic Description – ISBD* e *Anglo-American Cataloguing Rules – AACR*) e em normas reguladoras para a descrição (esquemas de codificação) dos campos de valor. Alves, Simionato e Santos (2012, p. 6) reforçam que

O padrão MARC 21 apresenta um esquema de metadados exaustivo e altamente estruturado, baseado em princípios, normas e códigos do domínio. [...] seu esquema apresenta um alto nível de detalhamento na descrição, necessita de esquemas externos para a construção de representações padronizadas e refinadas [...].

Embora seja amplamente utilizado no ambiente de bibliotecas, o padrão MARC 21 apresenta dificuldades quanto ao seu uso no ambiente Web, principalmente no que se refere à sintaxe necessária para compartilhar seus registros e promover a interoperabilidade dos mesmos, devido a estrutura de seu esquema de metadados. Para atender aos avanços tecnológicos e às necessidades que surgiram com a criação de ambientes digitais, a *Library of Congress* desenvolveu a versão MARCXML, que possui sua sintaxe em XML, permitindo que seu esquema de metadados se torne flexível e extensível, sem perder a especificidade inerente ao domínio bibliográfico.

Alves e Santos (2013, p. 93) afirmam que

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

Entretanto, o MARCXML também não atende completamente a necessidade de interoperabilidade dos dados na Web, além disso, o formato MARC 21 permaneceu inalterado em relação à sua sintaxe para intercâmbio e interoperabilidade na Web. Assim, houve a

necessidade de desenvolver um novo modelo de representação bibliográfica compatível com as tecnologias atuais e que pudesse ser compatível com as novas necessidades de compartilhamento de dados na Web, proporcionados pelo *Linked Data*.

3 Bibliographic Framework (BIBFRAME)

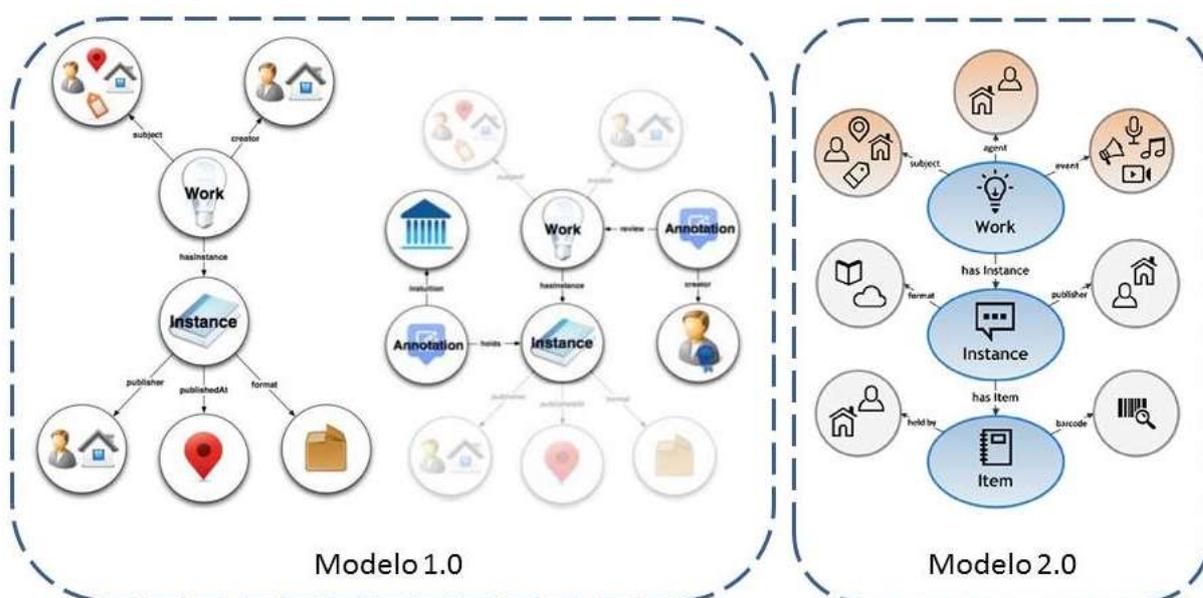
Segundo Silva (2013), o *Bibliographic Framework* (BIBFRAME) é um modelo de dados desenvolvido pela *Library of Congress* (LC) lançado oficialmente em maio de 2011, com o intuito de “[...] revisar e, a longo prazo, implementar um novo ambiente bibliográfico onde exista uma ‘network’ central que permita às bibliotecas um cenário de dados interligados.” (LIBRARY OF CONGRESS, 2012, p. 3, tradução nossa, grifo nosso). O modelo tem ainda o objetivo de “introduzir” a descrição bibliográfica como parte da *web* e dos dados interconectados, visando:

1. Diferenciar claramente o conteúdo conceitual e suas manifestações físicas;
2. Focar na identificação de forma inequívoca das entidades de informação;
3. Alavancar e expor as relações entre entidades.

(LIBRARY OF CONGRESS, 2012. p. 3, tradução nossa).

O BIBFRAME faz uso dos *Functional Requirements for Bibliographic Records* (FRBR), adaptando do modelo conceitual, porém não na íntegra, as entidades, atributos e relacionamentos. O BIBFRAME 1.0 consistia, na sua criação, em quatro classes principais: *Work*, *Instance*, *Authority* e *Annotation*. Entretanto, em abril de 2016, a LC lançou o modelo BIBFRAME 2.0, organizando a informação em três níveis de abstração: *Work*, *Instance* e *Item*, conforme apresentado na figura 1.

Figura 1 - Representação das entidades principais do BIBFRAME 1.0 e 2.0



Fonte: Adaptado de *Library of Congress* (2012) e *Library of Congress* (2016)

Além das três classes principais, o BIBFRAME conta com diversas outras classes e subclasses, descritas no *BIBFRAME Vocabulary*. Tais classes possuem propriedades,

utilizadas para “[...] definir formalmente os atributos e relacionamentos e representar os metadados dos recursos.” (RAMALHO, 2016, p. 299).

O BIBFRAME baseia-se ainda na arquitetura *Resource Description Framework* (RDF) e no uso de *links* e identificadores controlados, permitindo a inclusão das descrições no ambiente *web* e a inserção dos dados no *Linked Data* (SILVA et al., 2017).

O BIBFRAME ainda está em fase de testes e adaptação, portanto, ainda não foi implementado de forma definitiva em nenhum local (até o momento deste estudo). Entretanto, a LC disponibiliza algumas ferramentas em seu *site* para disseminar o uso do novo modelo, bem como o *BIBFRAME Editor* e o *software* de transformação de registros em MARCXML para BIBFRAME.

Algumas unidades de informação disponibilizam dados transformados de seus registros em BIBFRAME, como a *British Library* (Reino Unido), *Deutsche National Bibliothek* (Alemanha), *George Washington University Library* (EUA), *National Library of Medicine* (EUA) e *Princeton Library* (EUA) (ARAKAKI, 2016).

CONSIDERAÇÕES

A evolução dos instrumentos de representação no domínio bibliográfico acarretará numa nova transição para os catálogos das bibliotecas. Baseados principalmente nos modelos conceituais (FRBR e FRAD), as diretrizes RDA e os conceitos da *Web Semântica*, o processo de catalogação e os catálogos passarão por algumas modificações. Essas mudanças são fundamentais para maximizar a busca e recuperação das informações nos catálogos disponibilizados na *Web*.

Com o BIBFRAME, a forma de catalogação proposta pelo FRBR é mais visível, pois, a catalogação poderá ser fragmentada e separada em blocos, ao invés de um registro único como era proposto pela estrutura do MARC 21. Nesse sentido, uma vez feita a descrição da obra de um recurso informacional, basta fazer a ligação com o registro da manifestação, agilizando assim, o processo de catalogação dos recursos informacionais subsequentes.

Essa concepção é possível a partir da proposta do *Linked Data*, que consiste em ligar recursos e entidades a partir de *links*, facilitando a descrição dos recursos. Assim, as entidades (obra, manifestação, registros de autoridades, como pessoas, entidades coletivas, assuntos) poderão ser descritas separadamente e uma única vez, bastando relacionar uma ou mais entidades.

Entretanto, destaca-se que essa transição não ocorrerá rapidamente, pois, assim como as fichas catalográficas não acabaram completamente, o MARC também não acabará repentinamente. Portanto, existirá mais uma alternativa para descrição de recursos informacionais, principalmente para os que estarão no ambiente da *Web*.

REFERÊNCIAS

ALVES, Rachel Cristina Vesú; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. **Metadados no domínio bibliográfico**. Rio de Janeiro: Intertexto, 2013.

ALVES, Rachel Cristina Vesú; SIMIONATO, Ana Carolina; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Aspectos de granularidade na representação da informação no universo

bibliográfico. IN: ENCONTRO NACIONAL DE CATALOGADORES, 1., 2012, Rio de Janeiro. **Anais...** Rio de Janeiro: [S.n.], 2012. Disponível em: <<http://gepcat.blogspot.com.br/2012/10/i-enacat-e-iii-eepec-trabalhos.html>>. Acesso em: 21 dez. 2014.

ARAKAKI, Felipe Augusto. **Linked Data**: ligação de dados bibliográficos. 2016. 144 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Universidade Estadual Paulista, Marília, 2017. Disponível em: <<https://repositorio.unesp.br/handle/11449/147979>>. Acesso em: 25 jun. 2017.

CATARINO, Maria Elisabete; SOUZA, Terezinha Batista de. A representação descritiva no contexto da web semântica. **TransInformação**, Campinas, v. 24, n. 2, p. 77-90, maio/ago. 2012.

LIBRARY OF CONGRESS. **Bibliographic Framework as a web of data**: Linked Data model and supporting services. Washington, DC: Library of Congress, 2012.

LIBRARY OF CONGRESS. **Overview of the BIBFRAME 2.0 Model**. Washington, DC, 2016. Disponível em: <<http://www.loc.gov/bibframe/docs/bibframe2-model.html>>. Acesso em: 15 fev. 2017.

RAMALHO, Rogério Aparecido Sá. BIBFRAME: modelo de dados interligados para bibliotecas. **Informação & Informação**, Londrina, v. 21, n. 2, p. 292-306, maio/ago. 2016.

SILVA, Renata Eleutério da. **As tecnologias da Web Semântica no domínio bibliográfico**. 2013. 134 f. Dissertação (Mestrado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2013.

SILVA, Luciana Candida da et al. O código RDA e a iniciativa BIBFRAME: tendências da representação da informação no domínio bibliográfico. **Em Questão**, Porto Alegre, Online First, 2017. Disponível em: <<http://seer.ufrgs.br/index.php/EmQuestao/article/view/69549>>. Acesso em: 15 de jul. 2017.