

# Representação no domínio bibliográfico: um olhar sobre os Formatos MARC 21<sup>1</sup>

**Fabício Silva Assumpção**

**Bacharel em Biblioteconomia e mestre em Ciência da Informação pela UNESP. Doutorando do Programa de Pós-Graduação em Ciência da Informação da UNESP, Campus de Marília**

**Plácida Leopoldina Ventura Amorim da Costa Santos**

**Livre-Docente em Catalogação. Docente do Programa de Pós-Graduação em Ciência da Informação da UNESP, Campus de Marília.**

**<http://dx.doi.org/10.1590/1981-5344/2054>**

*A representação da informação tem sido realizada em diversos domínios, um deles o domínio bibliográfico, onde está pautada em instrumentos, princípios, modelos e tecnologias, tais como os padrões de metadados e as codificações. Entre os padrões de metadados utilizados nesse domínio, estão os Formatos MARC 21, com origens na década de 1960. Considerando o amplo uso desses padrões, tem-se por objetivo (1) destacar os propósitos que conduziram à criação dos Formatos MARC 21, (2) apresentar as características da codificação dos registros nesses formatos e (3) discutir as principais críticas sobre sua situação atual. Para tanto, é realizada uma revisão de literatura sobre a origem do MARC e seu desenvolvimento até o MARC 21 e sobre a codificação de registros. Assim, é apresentada a codificação com a XML e o esquema MARCXML, bem como críticas aos Formatos MARC 21. Conclui-se que, apesar das críticas, os Formatos MARC 21 continuam sendo amplamente utilizados e disseminados e que, apesar das vantagens oferecidas pela XML, a codificação com a norma ISO 2709, criada para o intercâmbio de dados em fitas magnéticas na década de 1960, ainda é utilizada.*

**Palavras-chave:** Formatos MARC 21; MARCXML; Padrões de metadados; Catalogação descritiva.

---

<sup>1</sup> Pesquisa financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## Representation in the bibliographic domain: a look at MARC 21 Formats

*Information representation has been done in several domains, one of which is the bibliographic domain, where information representation is based on instruments, principles, models and technologies, such as metadata standards and encodings. Among the standards used in the bibliographic domain, there are MARC 21 Formats that were created in 1960s. Considering the widespread use of these standards, we aim to (1) highlight the purposes that led to the creation of MARC 21 Formats, (2) present the features of record encoding in these standards and (3) discuss the main criticism about their current situation. To achieve these goals, we conducted a literature review about the origins of MARC and its development for MARC 21 and also about records encoding. Thus, we present the encoding with the XML and the MARCXML schema, as well as the criticism about MARC 21 Formats. We conclude that, despite the criticism, MARC 21 Formats are still widespread used and disseminated. We also conclude that, despite the advantages offered by XML, the encoding with ISO 2709, created for data interchange by using magnetic tape in 1960s, is still used.*

**Keywords:** MARC 21 Formats; MARCXML; Metadata standards; Descriptive cataloging.

Recebido em 30.03.2014 Aceito em 24.02.2015

### 1 Introdução

Inerente aos mais distintos sistemas de informação, a atividade de representação da informação tem sido realizada em vários domínios, um deles o domínio bibliográfico. Para as atividades de catalogação descritiva desenvolvidas nesse domínio são utilizados diversos instrumentos, construídos sobre princípios, modelos e aspectos teóricos e práticos e apoiados em diversas tecnologias.

Na literatura são encontradas categorizações para os instrumentos, princípios, modelos e tecnologias relacionadas à representação da informação tal como estudada e desenvolvida pela catalogação descritiva no domínio bibliográfico. Uma dessas categorizações é a de Picco e Ortiz Repiso (2012), apresentada na Figura 1.

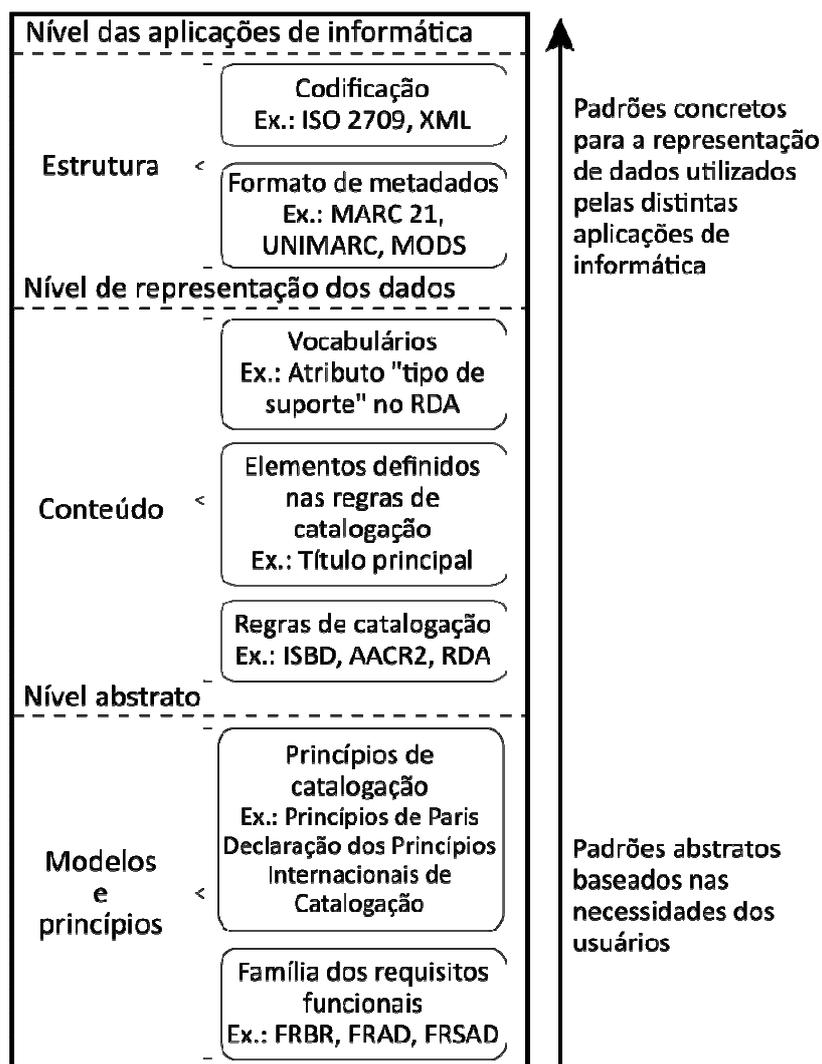


Figura 1 – Categorização dos instrumentos, princípios, modelos e tecnologias

Fonte: PICCO; ORTIZ REPISO (2012, p. 149, tradução nossa).

As autoras consideram três principais níveis: um nível abstrato abrangendo os modelos e princípios do domínio bibliográfico; um nível de representação dos dados, que compreende os instrumentos destinados ao conteúdo das representações; e um nível relacionado às aplicações de informática, compreendendo os formatos ou padrões de metadados e as codificações.

Cabe notar que, conforme a Figura 1, tanto as codificações quanto os padrões de metadados podem ser entendidos como estruturas. As codificações são estruturas que permitem o **processamento** dos registros catalográficos por aplicações de informática. Os padrões de metadados são estruturas que permitem a **identificação** dos dados nos registros catalográficos durante seu processamento. Assim, dadas essas possibilidades de interpretação da palavra estrutura, neste estudo são

utilizados, sempre que possível, os termos codificação e padrão de metadados, ao invés de estrutura.

Impulsionado pelos avanços das tecnologias de informática, o cenário internacional da catalogação descritiva atualmente contempla discussões sobre os elementos pontuados por Picco e Ortiz Repiso (2012). Os modelos conceituais e a Declaração dos Princípios Internacionais de Catalogação, influenciando os objetivos, a estrutura, as diretrizes e as instruções de novos conjuntos de regras de catalogação, fomentam questionamentos sobre os padrões de metadados e as codificações atualmente utilizadas.

Alguns dos padrões de metadados internacionalmente utilizados no domínio bibliográfico fazem parte dos Formatos MARC 21, conjunto de padrões de metadados que tem sua origem na década de 1960. Considerando o amplo uso desses padrões de metadados no domínio bibliográfico, tem-se por objetivo (1) destacar os propósitos que conduziram à criação dos Formatos MARC 21, (2) apresentar as características da codificação dos registros nesses formatos e (3) discutir as principais críticas sobre sua situação atual.

## 2 Das fitas magnéticas...

No domínio bibliográfico, as questões sobre a representação de recursos informacionais no ambiente digital estavam em debate já na década de 1960, mesmo que tal ambiente fosse pensado apenas como ferramenta para potencializar as atividades dos ambientes analógicos.

Em 1964, o *Council on Library Resources* (CLR) financiou um estudo dos possíveis métodos para a conversão de fichas catalográficas da *Library of Congress* (LC) para uma forma legível por máquina (*machine-readable*) com o objetivo de utilizar computadores na impressão de bibliografias. O estudo deu origem a um relatório que levou a realização da *First Conference on Machine-Readable Catalog Copy* em janeiro de 1965 (AVRAM, 1968, p. 3).

Dentre os tópicos de discussão dessa conferência estava o formato para registros catalográficos legíveis por máquina, os elementos bibliográficos compreendidos por esse formato, a distribuição de dados catalográficos legíveis por máquina e o futuro dos catálogos em fichas e impressos (AVRAM, 1968, p. 3). Entre as conclusões da conferência estavam:

3. A Library of Congress deve provavelmente incluir em seu registro legível por máquina todos os dados impressos na ficha catalográfica e informações adicionais. A maioria dos conferencistas foi favorável à codificação de tantos dados quanto possível para assegurar uma máxima recuperação no futuro. Parece ser desejável para a Library of Congress ir adiante com suas próprias necessidades e que outras bibliotecas utilizem o que quiserem de um registro legível por máquina da LC.

4. O registro legível por máquina seria utilizado para uma variedade de produtos bibliográficos, tais como catálogos em fichas, catálogos impressos, bibliografias, listas de aquisição, etc. (AVRAM, 1968, p. 3, tradução nossa).

A discussão desses tópicos, a realização de análises dos dados catalográficos na LC e a revisão do formato proposto conduziram a uma segunda conferência em novembro de 1965. Em fevereiro de 1966 foi realizada a terceira conferência, marcando o início oficial do *MACHine-Readable Cataloging (MARC) Pilot Project*. Em novembro do mesmo ano teve início o serviço semanal de distribuição de registros que, ao ser finalizado em junho 1968, havia distribuído aproximadamente 50 mil registros em fitas magnéticas (AVRAM, 1968, p. 4-7).

Durante o projeto piloto foi utilizado o formato MARC I. Com base na experiência obtida nesse projeto, foi desenvolvido o MARC II, que passou a ser utilizado a partir de março de 1969 na distribuição de registros de publicações monográficas em língua inglesa (AVRAM, 2003, p. 1714).

García Melero e García Camarero (1999 *apud* EITO BRUN, 2008, p. 148, tradução nossa) sintetizam os objetivos básicos que levaram a criação do MARC:

1. A criação de descrições bibliográficas reutilizáveis em distintos processos (aquisição, catalogação, recuperação e circulação ou empréstimo).
2. Normalizar a estrutura dos registros bibliográficos para facilitar seu intercâmbio entre bibliotecas.
3. Alcançar a independência dos dados bibliográficos frente aos sistemas de informação utilizados para seu tratamento automatizado.

Tendo como base o MARC II, formatos com características semelhantes surgiram em diversos locais, como o CAN/MARC (Canadá), o MONOCLE (França), o FINMARC (Finlândia), o IBERMARC (Espanha), o ANNAMARC (Itália) e o MARCAL (América Latina) (EITO BRUN, 2008, p. 147; MORENO; BRASCHER, 2007, p. 15), além do UNIMARC, desenvolvido e mantido pela *International Federation of Library Associations and Institutions* (IFLA).

No Brasil destaca-se a iniciativa da bibliotecária Alice Príncipe Barbosa que, em 1972, defendeu em seu mestrado o Formato Catalogação Legível por Computador (CALCO), baseado no Formato MARC II. Apesar da iniciativa, o CALCO não foi atualizado e caiu em desuso (SANTOS; CORRÊA, 2009, p. 49).

A importância dos Formatos MARC é claramente notada no intercâmbio de registros bibliográficos com informações de forma

padronizada, no planejamento e na implantação da catalogação cooperativa para redução de custos e retrabalhos.

Na década de 1980 o MARC II tornou-se USMARC (LIBRARY OF CONGRESS, 2006). Na década seguinte surgiu a denominação MARC 21 a partir da harmonização entre os Formatos USMARC e CAN/MARC, que passaram a ser publicado sob o novo nome: MARC 21.

O MARC 21 não é um novo formato. De 1994 a 1997 as comunidades de usuário do USMARC e do CAN/MARC trabalharam para eliminar todas as diferenças remanescentes em seus dois formatos já similares. A compatibilidade tem sido uma característica do processo de desenvolvimento de ambos os formatos por muitos anos. Em 1997 e no início de 1998, atualizações dos formatos foram publicadas para tornar as especificações idênticas. MARC 21, a continuação do USMARC e do CAN/MARC, publica os formatos em uma edição sob um novo nome. (LIBRARY OF CONGRESS, 1998, tradução nossa).

Atualmente, os Formatos MARC 21 compreendem cinco formatos: para dados bibliográficos, de autoridade, de coleção, de classificação e para informação comunitária.

Um registro MARC 21 envolve três componentes: a estrutura do registro, a designação do conteúdo e os dados de conteúdo (MACHINE-READABLE BIBLIOGRAPHIC INFORMATION COMMITTEE; NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE, 1996). Seguindo a categorização de Picco e Ortiz Repiso (2012), a estrutura do registro é dada pela codificação, a designação do conteúdo é provida pelo padrão de metadados e os dados de conteúdo estão de acordo com as regras de catalogação, seus elementos e vocabulários.

O formato legível por máquina desenvolvido na LC, na década de 1960, visava à transmissão de registros em fitas magnéticas, o que justifica o caráter sequencial da codificação utilizada originalmente na transmissão de registros MARC. Essa codificação foi reconhecida pelo *American National Standards Institute* (ANSI) em 1971 como a norma ANSI Z39.2-1971, e, em 1981, pela *International Organization for Standardization* (ISO) como a ISO 2709-1981, *Documentation - Format for bibliographic information interchange on magnetic tape* (BURGER, 1985, p. 73). A Figura 2 apresenta uma ilustração didática de um registro no Formato MARC 21 para Dados Bibliográficos codificado com a ISO 2709.

```
00473cam a2200169 a 45000010009000000050017000090
08004100026020001800067082001600085100002300101245
00450012425000120016926000350018130000290021665000
26002456510032002711592244920100930102920.0090925s
2009 nyua 000 1 eng d a9780385504225
00 a813/.54 222 1 aBrown, Dan, d1964- 14 aThe lost
symbol : ba novel / cDan Brown. a1st ed. aNew
York : bDoubleday, cc2009. a509 p. : bill. ; c25
cm. 0 aFreemasonry vFiction. 0 aWashington (D.C.)
vFiction.
```

Figura 2 – Registro codificado com a ISO 2709

Fonte: Adaptado de: <<http://lccn.loc.gov/2009464840>>. Acesso em: 9 fev. 2014.

O registro mostrado na Figura 2 constitui-se de uma linha contínua de caracteres, no entanto, é aqui apresentado com quebras de linhas apenas para possibilitar uma melhor exibição. Esse registro sequencial é intercambiado entre sistemas e então processado de acordo com o líder, com o diretório e com as etiquetas, indicadores e códigos de subcampo definidos no padrão de metadados Formato MARC 21 para Dados Bibliográficos (LIBRARY OF CONGRESS, 2013).

Apesar das normas para a codificação terem sido atualizadas no decorrer dos anos, nota-se que não houve mudanças significativas na codificação, de modo que os registros nos atuais Formatos MARC 21 são, em sua maior parte, codificados quase que da mesma forma com que eram codificados os registros na década de 1960, seja para propósitos de recuperação, por exemplo, via protocolo Z39.50, de importação entre sistemas de gerenciamento de bibliotecas ou de armazenamento em bancos de dados.

### 3 ... à marcação utilizando a XML

Os esforços da LC para utilizar registros nos Formatos MARC codificados de forma mais condizente com as tecnologias de informática de cada época iniciaram-se na década de 1990 com estudos sobre a utilização da *Standard Generalized Markup Language* (SGML).

A SGML é uma linguagem criada para facilitar a produção e o compartilhamento de documentos eletrônicos (MILLER; CLARKE, 2004, p. 2) ou, segundo Moreno e Brascher (2007, p. 15), “uma linguagem para descrever a estrutura dos documentos, ou qualquer tipo de dado textual, deixando a interpretação dos dados para outros programas”.

Os estudos desenvolvidos pela LC culminaram na criação de *Document Type Definitions* (DTD) para a codificação (marcação) de registros nos Formatos MARC utilizando a SGML. As DTDs especificam os

elementos de marcação que podem ocorrer em um documento, por exemplo, quais são os elementos permitidos, se são repetíveis ou não, em quais locais do documento podem ocorrer, etc., por isso são utilizadas para a validação dos documentos. A validação é o processo de checagem se a marcação presente no documento está de acordo com as especificações da DTD.

As DTDs SGML definiam um elemento SGML para cada campo e subcampo dos Formatos MARC. Por exemplo, para o campo 245 foram criados os elementos <mrcb245>, <mrcb245-a>, <mrcb245-b>, etc. Esses elementos eram agrupados em outros elementos SGML como <mrcb-title-and-title-related>, <mrcb-subject-access>, etc. Para os campos de controle foram especificados elementos para cada posição de caractere. Essa abordagem utilizada na codificação de registros nos Formatos MARC utilizando a SGML resultou em DTDs extremamente largas e complexas (EITO BRUN, 2008, p. 153).

Posteriormente, com o lançamento da *Extensible Markup Language* (XML) pelo *World Wide Web Consortium* (W3C) em 1998, a LC voltou seus esforços para essa nova possibilidade de codificação. Foram criadas, então, duas DTDs para a codificação de registros MARC 21 utilizando a XML. A primeira DTD era utilizada para registros bibliográficos, de itens e de informação comunitária; a segunda destinava-se aos registros de autoridade e de classificação.

Essas DTDs XML definiam todos os elementos que poderiam aparecer em um registro MARC 21 codificado com a XML e especificavam como esses elementos seriam rotulados e representados nessa codificação (TAYLOR; JOUDREY, 2009, p. 153). A Figura 3 apresenta o fragmento de um registro no Formato MARC 21 para Dados Bibliográficos codificado de acordo com a DTD XML criada pela LC.

```
<mrcb-main-entry>
  <mrcb100 i1="i1-1" i2="i2-blank">
    <mrcb100-a>Brown, Dan,</mrcb100-a>
    <mrcb100-d>1964-</mrcb100-d>
  </mrcb100>
</mrcb-main-entry>
<mrcb-title-and-title-related>
  <mrcb245 i1="i1-1" i2="i2-4">
    <mrcb245-a>The lost symbol :</mrcb245-a>
    <mrcb245-b>a novel /</mrcb245-b>
    <mrcb245-c>Dan Brown.</mrcb245-c>
  </mrcb245>
</mrcb-title-and-title-related>
<mrcb-edition-imprint-etc>
  <mrcb250 i1="i1-blank" i2="i2-blank">
    <mrcb250-a>1st ed.</mrcb250-a>
  </mrcb250>
  <mrcb260 i1="i1-blank" i2="i2-blank">
    <mrcb260-a>New York :</mrcb260-a>
    <mrcb260-b>Doubleday,</mrcb260-b>
    <mrcb260-c>c2009.</mrcb260-c>
  </mrcb260>
</mrcb-edition-imprint-etc>
<mrcb-physical-description>
  <mrcb300 i1="i1-blank" i2="i2-blank">
    <mrcb300-a>509 p. :</mrcb300-a>
    <mrcb300-b>ill. ;</mrcb300-b>
    <mrcb300-c>25 cm.</mrcb300-c>
  </mrcb300>
</mrcb-physical-description>
```

Figura 3 – Fragmento de um registro codificado com a DTD XML

Fonte: Dados da pesquisa.

Embora as DTDs criadas pela LC fossem destinadas à codificação com a XML, ainda mantinham as características das DTDs SGML, tais como o fato de serem documentos demasiado extensos. Segundo Siqueira (2003, p. 83), se impressas, as duas DTDs XML corresponderiam a um documento com mais de 600 páginas.

Para Eito Brun (2008, p. 152), a utilização da XML para a codificação de registros MARC traria como vantagens a utilização de uma codificação universalmente aceita, mais simples que a provida pela ISO 2709 e de processamento mais fácil, o que resultaria em uma menor complexidade no desenvolvimento de aplicações informáticas para o processamento de registros MARC.

Eito Brun (2008, p. 152) aponta também que um dos objetivos que levaram à codificação de registros utilizando a XML foi evitar a segregação do MARC frente a outros padrões de metadados projetados para a Web e que prometiam tornarem-se grandes protagonistas no ambiente das bibliotecas digitais. Ainda segundo o autor, manter a codificação tradicional do MARC como o único meio para a transferência de registros influenciaria negativamente nas possibilidades de intercâmbio de informações e na exposição das descrições a outros serviços de informação.

Originalmente, DTDs foram criadas para validar documentos SGML, no entanto foram utilizadas também para a validação de documentos XML, embora ainda preservassem características da SGML. Como uma alternativa ao uso de DTDs para validar documentos XML, o W3C aprovou em 2001 a linguagem *XML Schema*, utilizada na criação de documentos que desempenham funções semelhantes às desempenhadas por uma DTD, no entanto, projetada para a e seguindo a XML.

Segundo o W3C (2012, tradução nossa), os “esquemas XML expressam vocabulários<sup>2</sup> compartilhados e permitem que máquinas executem regras feitas por pessoas. Eles provêm meios para definir a estrutura, o conteúdo e a semântica de documentos XML”.

Utilizando-se da *XML Schema*, em 2002, a LC substituiu as DTDs XML do MARC pelo *MARC 21 XML Schema*, mais conhecido como MARCXML (LIBRARY OF CONGRESS, 2012b).

No desenvolvimento do MARCXML foi adotada uma abordagem diferente daquela utilizada nas DTDs. Ao invés de criar um elemento para cada campo e subcampo, foram criados elementos XML para os diferentes tipos de elementos do MARC 21 (*leader*, *controlfield*, *datafield* e *subfield*), sendo que as *tags* dos campos, os indicadores e os códigos de subcampos definidos nos Formatos MARC 21 seriam indicados por meio de atributos XML. O resultado dessa abordagem foi um esquema único e mais simples para todos os tipos de registros MARC 21 (bibliográfico, de autoridade, etc.) (EITO BRUN, 2008, p. 154).

A Figura 4 traz um registro no Formato MARC 21 para Dados Bibliográficos codificado em XML de acordo com o MARCXML.

```
<marc:datafield tag="100" ind1="1" ind2="">
  <marc:subfield code="a">Brown, Dan,</marc:subfield>
  <marc:subfield code="d">1964-</marc:subfield>
</marc:datafield>
<marc:datafield tag="245" ind1="1" ind2="4">
  <marc:subfield code="a">The lost symbol :</marc:subfield>
  <marc:subfield code="b">a novel /</marc:subfield>
  <marc:subfield code="c">Dan Brown.</marc:subfield>
</marc:datafield>
<marc:datafield tag="250" ind1="" ind2="">
  <marc:subfield code="a">1st ed.</marc:subfield>
</marc:datafield>
<marc:datafield tag="260" ind1="" ind2="">
  <marc:subfield code="a">New York :</marc:subfield>
  <marc:subfield code="b">Doubleday,</marc:subfield>
  <marc:subfield code="c">c2009.</marc:subfield>
</marc:datafield>
<marc:datafield tag="300" ind1="" ind2="">
  <marc:subfield code="a">509 p. :</marc:subfield>
  <marc:subfield code="b">ill. ;</marc:subfield>
  <marc:subfield code="c">25 cm.</marc:subfield>
</marc:datafield>
```

Figura 4 – Registro MARC 21 codificado com o MARCXML

<sup>2</sup> Neste estudo, um vocabulário, como descrito pelo W3C, é entendido como um padrão de metadados.

Fonte: Adaptado de: <<http://lccn.loc.gov/2009464840>>. Acesso em: 9 fev. 2014.

As Figuras 2, 3 e 4 apresentam o mesmo registro no Formato MARC 21 para Dados Bibliográficos, no entanto, o fazem utilizando diferentes codificações. O conteúdo do registro e a designação do conteúdo (as etiquetas dos campos, os indicadores e os códigos dos subcampos) permanecem os mesmos, explicitando que, no caso dos Formatos MARC 21, a codificação não interfere no padrão de metadados e nem no conteúdo dos registros catalográficos.

Para Eito Brun (2008, p. 155), o MARCXML tem por objetivos:

- a) facilitar a codificação de registros nos Formatos MARC 21 utilizando a XML (mediante um processo de conversão);
- b) facilitar a captura e a integração de registros MARC nos processos de coleta de metadados (*metadata harvesting*), característicos dos projetos *Open Archives Initiative* (OAI) e similares;
- c) permitir a descrição original de recursos utilizando os Formatos MARC 21 com a XML, ou seja, permitir que registros MARC sejam criados diretamente como documentos XML usando uma ferramenta de edição adequada;
- d) facilitar a inclusão de dados de registros MARC 21 em textos completos também codificados com a XML, o que é uma possibilidade oferecida por padrões como o TEI P5, da *Text Encoding Initiative*, e permitida mediante o uso de *namespaces* (espaços nominais) XML.

A facilidade, do ponto de vista de um profissional da Ciência da Computação, em lidar com registros nos Formatos MARC codificados com a XML é abordada por Siqueira (2003). Segundo o autor, a forma de representação de um registro MARC 21 em XML torna mais simples o desenvolvimento de programas para a importação de dados nesse formato, principalmente considerando que: (1) a XML é de conhecimento dos programadores; (2) as ferramentas existentes trazem recursos que facilitam a recuperação das *tags*, posições, indicadores, subcampos e seus respectivos valores; (3) a própria estrutura da XML provê um meio de validação dos dados; e (4) um registro MARC 21 no formato XML pode ser apresentado de várias formas por meios de folhas de estilo de transformação.

Em síntese, a tradicional codificação de registros nos Formatos MARC 21, que ocorre no nível da estrutura do registro, segue as especificações contidas nas normas ISO 2709 e ANSI Z39.2, enquanto que a codificação com a SGML segue as DTDs SGML e a codificação com a XML pode seguir tanto as DTDs XML do MARC 21 quanto o MARCXML, como mostrado na Figura 5.

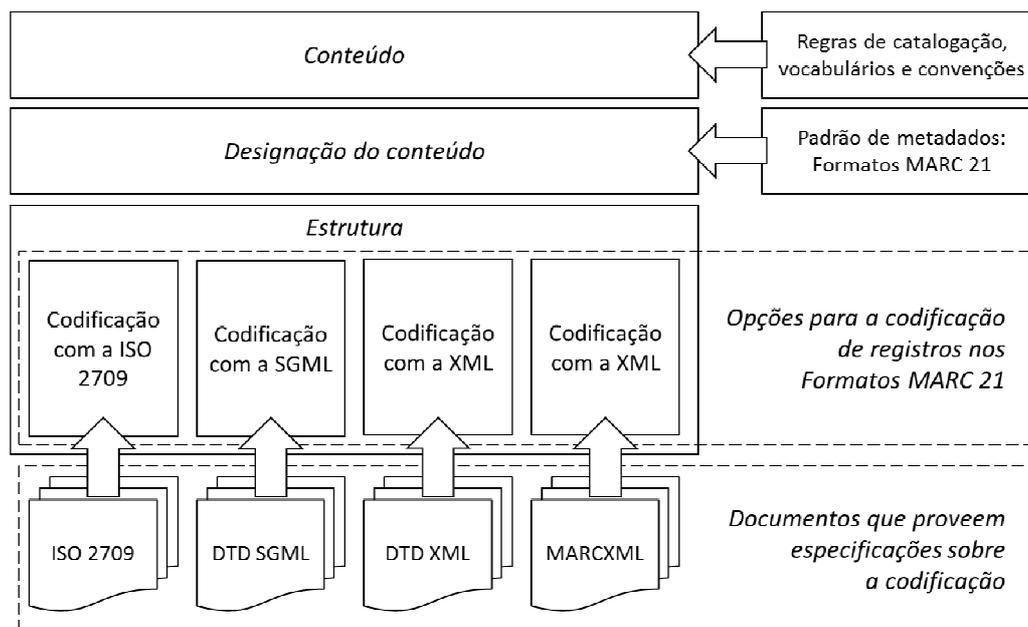


Figura 5 – Codificação de registros nos Formatos MARC 21 desenvolvidas pela LC

Fonte: Dados da pesquisa.

Nota-se também que para a codificação de registros MARC com a XML foram conduzidos outros projetos além dos da LC, como é o caso do XMLMARC realizado pela *Lane Medical Library* da Universidade de Stanford, nos Estados Unidos, e o BiblioML, subsidiado pelo Ministério da Cultura e Comunicação da França (EITO BRUN, 2008, p. 161).

Cabe observar que a codificação especificada pela ISO 2709 e pela ANSI Z39.2 teve sua origem junto com o padrão de metadados MARC, o que fez com que tal padrão e os padrões que dele derivaram (MARC II, UNIMARC, MARC 21, etc.) se tornassem bastante dependentes dessa codificação.

Como apresentado nesta seção, além da codificação tradicional sequencial, tem-se a possibilidade de codificar registros nos Formatos MARC 21 utilizando a XML, que está em maior consonância com as tecnologias de informática atuais. Mesmo com essa possibilidade, nota-se que a comunidade de bibliotecas encontra-se ainda presa à codificação da ISO 2709, utilizando uma estrutura sequencial desenvolvida para fitas magnéticas. Esse, dentre outros elementos, constitui-se como pontos de crítica aos Formatos MARC 21 e à sua utilização, como mostrado no item seguinte.

## 4 Críticas aos Formatos MARC

Diversas críticas aos Formatos MARC 21 são encontradas na literatura. Um ponto frequentemente destacado pelos autores é que, em razão dos Formatos MARC terem sido criados na década de 1960 para

possibilitar a produção de fichas catalográficas, um registro MARC 21, ainda hoje, simula uma ficha catalográfica no ambiente digital.

Segundo Coyle (2004, p. 166), o registro MARC foi criado como um reflexo digital das regras de catalogação de seu tempo, regras estas que não são tão diferentes das regras de catalogação atuais. Essas regras originalmente foram projetadas para a produção de fichas e ainda mantêm essa característica, o que é possível observar por meio dos cabeçalhos das entradas principais, da forma invertida dos nomes e do agrupamento de elementos em segmentos semelhantes a parágrafos.

Danskin (2004, p. 115) acrescenta à discussão outros aspectos. Para o autor, o registro MARC moderno é uma "reencarnação" da ficha catalográfica e tão ineficiente quanto esta, pois, toda vez que um catalogador cria um registro para uma nova edição de uma obra que a biblioteca já possui um exemplar, ele digita (ou copia) dados que já estão presentes na base de dados, mas não estão estruturados de modo que possam ser reutilizados. Para Danskin, isso é ignorar a capacidade das bases de dados relacionais presentes na maioria dos modernos sistemas de gerenciamento de bibliotecas.

Mönch e Aalberg (2003, p. 405) descrevem algo semelhante. Os autores exemplificam que, uma busca por uma obra literária que existe em numerosas edições e traduções está suscetível a gerar um grande conjunto de resultados, pois cada edição ou tradução está representada por um registro individual que não está relacionado a outros registros que descrevem a mesma obra.

Segundo Gorman (1997, tradução nossa), "[...] o MARC tem suas origens na ficha catalográfica. Isso não é somente visto na ordem dos campos do MARC, que preserva exatamente a ordem dos dados na ficha catalográfica, incluindo a separação do cabeçalho da 'entrada principal' (1XX) de outros pontos de acesso (7XX)". Gorman (1997, tradução nossa) complementa que

Todos os problemas que temos com o MARC derivam da violação inicial da lei fundamental da automação de bibliotecas – "nunca automatize apenas o que você tem". Anos atrás, clamei por uma reforma completa do MARC que deveria, essencialmente, substituir por simples registros (nomes, descrições, assuntos) com muitas e complexas conexões o que temos hoje – complexos registros com poucas conexões. É evidente que meu chamado não foi ouvido. Como uma consequência, estamos lidando com os efeitos de milhões de registros MARC e centenas de sistemas baseados nesses registros sem a capacidade de tirar vantagem da sofisticação dos modernos sistemas online.

Para Serra (2013, p. 3):

Outra questão limitante do formato é relativa à dificuldade para representar hierarquias. O MARC proporciona uma representação horizontal, sem vínculos entre registros, dificultando a descrição bibliográfica, principalmente em um contexto digital, questão esta que é bem atendida por linguagens XML. O formato não permite a inclusão de mídia rica como capas ou sumários, limitando a aplicação de arquivos na tag 856, vinculando um conjunto de metadados a arquivos armazenados na web, servidores ou repositórios.

Do ponto de vista da construção de catálogos digitais, Fusco (2011) destaca os problemas resultantes da modelagem dos bancos de dados utilizando os Formatos MARC 21 como modelo. Para o autor:

É importante destacar que o MARC21 é um formato para intercâmbio de dados, que informa como um registro bibliográfico e catalográfico deve estar descrito para que sua importação ou exportação ocorra com sucesso, porém o modelo de dados do catálogo não necessariamente deve estar organizado estruturalmente no mesmo formato de um registro MARC21. (FUSCO, 2011, p. 57).

Fusco expõe que o uso indevido dos Formatos MARC 21 resulta em anomalias nos bancos de dados:

Isso [anomalias no banco de dados] ocorre porque padrões de metadados como o formato MARC têm uma estrutura linear em suas estruturas de registros, ou seja, um modelo bidimensional de atributo-valor e determinam o formato de entrada e de intercâmbio de dados, sem determinar o conceito de domínio dos objetos representados com seus relacionamentos. (FUSCO, 2011, p. 119).

Miller e Clarke (2003) apontam alguns problemas com os Formatos MARC 21, em síntese:

- a) inconsistência dos dados: um mesmo tipo de dado é registrado em diferentes campos/subcampos de diferentes formas (por exemplo, datas);
- b) redundância de dados: um mesmo dado é registrado em mais de um campo/subcampo, ora de modo codificado, ora literalmente (por exemplo, idioma);
- c) mistura de dados e seus atributos;
- d) dispersão e irregularidade dos relacionamentos;
- e) extrema complexidade na codificação.

Os autores acrescentam que:

Onde há informação hierárquica envolvida, a estrutura relativamente plana do MARC torna isso difícil de ser refletido. As limitações do tamanho do campo variam de sistema para sistema. A granularidade injustificada existe, enquanto que a funcionalidade indicando a granularidade está faltando. (MILLER; CLARKE, 2003, p. 116, tradução nossa).

Miller e Clarke (2003, p. 116) finalizam que os problemas são tantos que o MARC deve ser completamente reavaliado e reconstruído. Afirmam também que o MARC efetivamente evita que as bibliotecas tirem total vantagem da XML e das tecnologias relacionadas e coloca as bibliotecas em desvantagem na arena competitiva do gerenciamento informacional. "Apesar de sua vida longa e de suas contribuições úteis, o MARC representa agora mais uma desvantagem que uma vantagem" (MILLER; CLARKE, 2003, p. 116, tradução nossa).

Yee (2004, p. 165) destaca que mesmo os bem informados sobre MARC 21 reconhecem que esse padrão de metadados tem instalado um alicerce menor do que o instalado pela XML, e que em algum momento no futuro podemos nos planejar para uma migração de nossos dados para algo semelhante à XML ou para um de seus sucessores, afim de termos acesso a um amplo mercado de soluções de software e hardware para o problema do controle bibliográfico.

Eito Brun (2008, p. 151) destaca que o MARC constitui-se como o principal mecanismo para a codificação e o intercâmbio automatizado de registros em bibliotecas e que muitas das vantagens oferecidas pelo MARC também são oferecidas pela XML e SGML: capacidade de diferenciar os dados contidos em um registro, independência de fabricantes de software, facilidade para o processamento automatizado, etc. Para o autor, considerando que as linguagens de marcação SGML e XML são posteriores aos Formatos MARC e se baseiam em um modelo de registro mais moderno e facilmente processável, não devemos estranhar que tenham se apresentado como uma alternativa ao MARC.

Cabe notar que, por mais que Yee e Eito Brun mencionem que a XML é uma alternativa aos Formatos MARC 21 ou que a XML pode substituir tal padrão de metadados, entende-se aqui que a XML apresenta-se apenas como uma alternativa à codificação de registros MARC 21 (como é a proposta do MARCXML, descrita anteriormente). Entende-se também que a XML pode ser utilizada na codificação de registros criados com outros padrões de metadados, sendo que esses outros padrões, sim, podem substituir os Formatos MARC 21.

Sobre a codificação de registros nos Formatos MARC 21 utilizando a XML, Miller e Clarke (2003, p. 103, tradução nossa) destacam que:

Os recentes esforços para modernizar o MARC têm, em sua maioria, envolvido uma tradução literal e sem crítica dos campos, indicadores e subcampos para elementos e atributos XML. Infelizmente, isso incorpora no novo meio os problemas

inerentes ao MARC e não permite que se faça uso das reais vantagens da XML.

Eito Brun (2008, p. 151-152) apresenta uma visão semelhante. Para o autor, as iniciativas para a codificação de registros MARC com a XML, frequentemente chamadas de "x-emelificação do MARC", propõem o uso da XML para representar e transferir registros, no entanto, sem questionar a organização lógica do registro ou o significado dos campos, subcampos e indicadores.

Com o objetivo de explorar os vários problemas que alguns autores têm associado aos Formatos MARC 21, Yee (2004) analisa e define quatro tipos de problemas:

- a) problemas que na verdade não são culpa do MARC 21, mas sim das regras e das práticas de catalogação que proveem conteúdo para os registros MARC 21, por exemplo, o AACR2r;
- b) problemas que realmente não são problemas, ao invés disso, são soluções para problemas imperfeitamente compreendidos pelos autores;
- c) problemas devido ao ambiente de catalogação compartilhada para o qual o MARC 21 foi projetado;
- d) problemas causados ou parcialmente causados pelo MARC 21 e que talvez possam ser solucionados nos processos de migração dos dados para um novo padrão de estrutura de dados no futuro.

Sobre os problemas do quarto tipo, Yee (2004, p. 166, tradução nossa) destaca que uma parte deles poderia ser resolvida nos software e não no próprio Formato MARC 21.

Infelizmente, muitos dos problemas que são atribuídos ao MARC 21 são, na verdade, problemas que derivam do insucesso dos distribuidores de software em suportar as capacidades totais do MARC 21. Às vezes isso se deve às considerações de ordem financeira (o desenvolvimento é feito somente quando um número significativo de clientes será beneficiado); às vezes isso ocorre devido à falta de entendimento que os distribuidores de software têm do MARC 21, dos registros catalográficos, dos problemas que surgem em complexas bases de dados de registros bibliográficos e dos problemas de face pública no acesso aos catálogos online de acesso público. (YEE, 2004, p. 166, tradução nossa).

Como bem categorizados por Yee (2004), os problemas associados aos Formatos MARC 21 são de naturezas diversas. A inadequação da codificação sequencial às atuais tecnologias de informática se constituiu

como uma real crítica aos Formatos MARC 21, no entanto, para esse problema, uma solução foi apresentada há mais de dez anos com o surgimento do MARCXML.

Na codificação de registros MARC 21 com a XML, uma mudança apenas para uma forma de codificação mais recente, em maior consonância com as atuais tecnologias de informática, não é suficiente para garantir a adequação dos Formatos MARC 21 ao cenário em que o domínio bibliográfico se encontra no começo do século XXI.

## 5 Considerações finais

Apesar das críticas, os Formatos MARC 21 continuam sendo amplamente utilizados e disseminados para o intercâmbio de dados catalográficos em ambientes digitais. Apesar das vantagens oferecidas pela codificação com a XML, inclusive no desenvolvimento de software para o processamento de registros MARC 21, ainda persiste, na maior parte dos casos, a codificação utilizando a estrutura sequencial definida pelas normas ISO 2709 e ANSI Z39.2.

Junto aos esforços para a utilização da XML na codificação de registros MARC 21, a LC tem projetado padrões de metadados que apresentam alternativas ao uso dos tradicionais formatos. Dentre esses padrões estão o *Metadata Object Description Schema* (MODS) (Esquema de Metadados para a Descrição de Objeto) e o *Metadata Authority Description Schema* (MADS) (Esquema de Metadados para a Descrição de Autoridade), ambos criados para o uso com a XML e especificados por esquemas XML.

O MODS e o MADS apresentam grande compatibilidade com os tradicionais Formatos MARC 21, embora, de modo geral, não possibilitem o registro de dados com o mesmo nível de especificidade dado pelos Formatos MARC 21 (ASSUMPÇÃO; SANTOS, 2013).

O MODS, devido a sua alta compatibilidade com o Formato MARC 21 para Dados Bibliográficos, pode ser escolhido por bibliotecas como um padrão de metadados para a descrição de recursos informacionais, apresentando uma alternativa ao padrão Dublin Core<sup>3</sup> (TAYLOR; JOUDREY, 2009, p. 153).

Apesar de mais de uma década de existência do MODS e do MADS, observa-se, principalmente na literatura nacional, a falta de estudos e de relatos sobre a implantação desses padrões. Ao passo que é possível encontrar diversos estudos sobre a utilização, no domínio bibliográfico, de um padrão não projetado para atender as necessidades de tal domínio, que é o caso do padrão Dublin Core.

Em maio de 2011 a LC fez um anúncio sobre sua iniciativa para uma nova estrutura bibliográfica. Afirmando a inadequação dos Formatos MARC 21 às necessidades atuais, inclusive àquelas relacionadas à utilização do *Resource Description and Access* (RDA), a LC propôs uma iniciativa cuja

---

<sup>3</sup> O padrão de metadados Dublin Core surgiu em 1995 com o objetivo possibilitar a criação de descrições amplas e gerais de recursos informacionais.

atividade central é o desenvolvimento de um novo meio para a captura e o compartilhamento de dados bibliográficos e que incluirá a busca por um substituto para os Formatos MARC 21 (LIBRARY OF CONGRESS, 2011).

A iniciativa da LC, atualmente apresentada no modelo e no vocabulário *Bibliographic Framework* (BIBFRAME), configura-se como base “para o futuro da descrição bibliográfica que ocorre sobre, na e como parte da Web e do mundo conectado em que vivemos” (LIBRARY OF CONGRESS, 2012a, p. 3, tradução nossa), incorporando questões sobre a utilização de tecnologias como o modelo de dados *Resource Description Framework* (RDF), além da já consolidada XML.

O BIBFRAME, apesar dos avanços já realizados em seu desenvolvimento, ainda encontra-se em fase inicial, sendo conduzidos os primeiros testes e realizados os ajustes necessários. Assim, entende-se que, para o intercâmbio de dados no domínio bibliográfico, os Formatos MARC 21 – apesar de suas críticas – ainda perdurarão como padrões de metadados amplamente utilizados pelos próximos anos.

Em tempos de um rápido desenvolvimento tecnológico com implicações diretas e indiretas nas atividades de representação de recursos informacionais e de intercâmbio de dados catalográficos, espera-se que este artigo, por apresentar um olhar sobre os Formatos MARC 21, sobre seu desenvolvimento e sobre suas codificações, tenha contribuído para a área da Ciência da Informação e, em especial, para as discussões acerca da representação de recursos informacionais no domínio bibliográfico.

## Referências

ASSUMPÇÃO, F. S.; SANTOS, P. L. V. A. da C. Metadata Authority Description Schema (MADS): uma alternativa à utilização do formato MARC 21 para dados de autoridade. *Informação & Informação*, v. 18, n. 1, p. 106-126, 2013. Disponível em: <<http://dx.doi.org/10.5433/1981-8920.2013v18n1p106>>. Acesso em: 9 fev. 2014.

AVRAM, H. D. *The MARC Pilot Project: final report on a project sponsored by The Council on Library Resources, Inc.* Washington, D.C.: Library of Congress, 1968. Disponível em: <<http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED029663>>. Acesso em: 9 fev. 2014.

AVRAM, H. D. Machine-Readable Cataloging (MARC) Program. *In: ENCYCLOPEDIA of Library and Information Science*. New York: Marcel Dekker, 2003. p. 1712-1730. Disponível em: <[http://polaris.gseis.ucla.edu/gleazer/260\\_readings/Avram.pdf](http://polaris.gseis.ucla.edu/gleazer/260_readings/Avram.pdf)>. Acesso em: 9 fev. 2014.

BURGER, R. H. *Authority work: the creation, use, maintenance, and evaluation of authority records and files*. Littleton: Libraries Unlimited, 1985.

COYLE, K. Future considerations: the functional library systems record. *Library Hi Tech*, n. 22, v. 2, p. 166-174, 2004. Disponível em: <<http://dx.doi.org/10.1108/07378830410524594>>. Acesso em: 9 fev. 2014.

DANSKIN, A. Mature consideration: developing bibliographic standards and maintaining values. *New Library World*, n. 105, v. 3/4, p. 113-117, 2004. Disponível em: <<http://dx.doi.org/10.1108/03074800410526730>>. Acesso em: 9 fev. 2014.

EITO BRUN, R. *Lenguajes de marcas para la gestión de recursos digitales: aproximación técnica, especificaciones y referencia*. Gijón: Trea, 2008.

FUSCO, E. *Aplicação dos FRBR na modelagem de catálogos bibliográficos digitais*. São Paulo: Cultura Acadêmica, 2011.

GARCÍA MELERO, L. Á.; GARCÍA CAMARERO, E. *Automatización de bibliotecas*. Madrid: Arco/Libros, 1999 *apud* EITO BRUN, R. *Lenguajes de marcas para la gestión de recursos digitales: aproximación técnica, especificaciones y referencia*. Gijón: Trea, 2008.

GORMAN, M. What is the future of cataloguing and cataloguers? *In: IFLA GENERAL CONFERENCE, 63., 1997, Copenhagen. Conference Programme and Proceedings*. [S.l.]: IFLA, 1997. Disponível em: <<http://archive.ifla.org/IV/ifla63/63gorm.htm>>. Acesso em: 9 fev. 2014.

LIBRARY OF CONGRESS. *Bibliographic Framework Initiative General Plan*. Washington, 2011. Disponível em: <<http://www.loc.gov/marc/transition/pdf/bibframework-10312011.pdf>>. Acesso em: 29 set. 2012.

LIBRARY OF CONGRESS. *Frequently Asked Questions (FAQ)*. Washington, 2006. Disponível em: <<http://www.loc.gov/marc/faq.html#definition>>. Acesso em: 9 fev. 2014.

LIBRARY OF CONGRESS. *MARC 21 Format for Bibliographic Data*. Washington, D.C., 2013. 1999 Edition. Update No. 1 (October 2001) through Update No. 17 (September 2013). Disponível em: <<http://www.loc.gov/marc/bibliographic/ecbdhome.html>>. Acesso em: 9 fev. 2014.

LIBRARY OF CONGRESS. *Bibliographic Framework as a Web of Data: Linked data model and supporting services*. Washington, 2012a. Disponível em: <<http://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>>. Acesso em: 9 fev. 2014.

LIBRARY OF CONGRESS. *MARC 21 XML Schema: official web site*. Washington, 2012b. Disponível em: <<http://www.loc.gov/standards/marcxml/>>. Acesso em: 9 fev. 2014.

LIBRARY OF CONGRESS. *MARC 21: Harmonized USMARC and CAN/MARC*. Washington, 1998. Disponível em: <<http://www.loc.gov/marc/anmarc21.html>>. Acesso em: 9 fev. 2014.

MACHINE-READABLE BIBLIOGRAPHIC INFORMATION COMMITTEE; NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE. *The MARC 21 Formats: background and principles*. Revised November 1996. Washington: Library of Congress, 1996. Disponível em: <<http://www.loc.gov/marc/96principl.html>>. Acesso em: 9 fev. 2014.

MILLER, D. R.; CLARKE, K. S. *Putting XML to work in the library: tools for improving access and management*. Chicago: American Library Association, 2004.

MÖNCH, C.; AALBERG, T. Automatic Conversion from MARC to FRBR. *Lecture Notes in Computer Science*, v. 2769, p. 405-411, 2003. Disponível em: <[http://dx.doi.org/10.1007/978-3-540-45175-4\\_37](http://dx.doi.org/10.1007/978-3-540-45175-4_37)>. Acesso em: 9 fev. 2014.

MORENO, F. P.; BRASCHER, M. MARC, MARCXML e FRBR: relações encontradas na literatura. *Informação & Sociedade: Estudos*, João Pessoa, v. 17, n. 3, p. 13-25, set./dez. 2007. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/834>>. Acesso em: 9 fev. 2014.

PICCO, P.; ORTIZ REPISO, V. RDA, el nuevo código de catalogación: cambios y desafíos para su aplicación. *Revista Española de Documentación Científica*, v. 35, n. 1, p. 145-173, enero-marzo 2012. Disponível em: <<http://dx.doi.org/10.3989/redc.2012.1.848>>. Acesso em: 9 fev. 2014.

SANTOS, P. L. V. A. C.; CORRÊA, R. M. R. *Catálogo: trajetória para um código internacional*. Niterói: Intertexto, 2009.

SERRA, L. G. O formato MARC e o RDA: tempos de mudanças? In: ENCONTRO INTERNACIONAL DE CATALOGADORES, 9., 2013, Rio de Janeiro; ENCONTRO NACIONAL DE CATALOGADORES, 2., 2013, Rio de Janeiro. *Anais...* Rio de Janeiro: Fundação Biblioteca Nacional, 2013. Disponível em: <<http://www.enacat.ufscar.br/index.php/eic-enacat/eic-enacat/paper/view/13>>. Acesso em: 9 fev. 2014.

SIQUEIRA, M. A. *XML na Ciência da Informação: uma análise do MARC 21*. 2003. 133 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003. Disponível em: <[http://www.athena.biblioteca.unesp.br/exlibris/bd/bma/33004110043P4/2003/siqueira\\_ma\\_me\\_mar.pdf](http://www.athena.biblioteca.unesp.br/exlibris/bd/bma/33004110043P4/2003/siqueira_ma_me_mar.pdf)>. Acesso em: 9 fev. 2014.

TAYLOR, A. G.; JOUDERY, D. N. *The organization of information*. 3rd ed. Westport: Libraries Unlimited, 2009.

W3C. *XML Schema*. Revision 1.153, date: 2012/06/22. Cambridge, 2012. Disponível em: <<http://www.w3.org/XML/Schema>>. Acesso em: 9 fev. 2014.

YEE, M. M. New perspectives on the shared cataloging environment and a MARC 21 shopping list. *Library Resources & Technical Services*, v. 48, n.

3, p. 165-178, 2004. Disponível em:  
<<http://escholarship.org/uc/item/6z76m6p9>>. Acesso em: 9 fev. 2014.