

REPRESENTAÇÃO E DESCRIÇÃO DE RECURSOS INFORMACIONAIS: aspectos estruturantes no delineamento de ambientes informacionais digitais¹

REPRESENTATION AND DESCRIPTION OF INFORMATION RESOURCES: structural aspects in the design of digital informational environments

Fabiano Ferreira de CASTRO

Doutorando e Mestre em Ciência da Informação, Universidade Estadual Paulista, Marília
e-mail: fabianofcastro@gmail.com

Plácida Leopoldina Ventura Amorim da Costa SANTOS

Professora Doutora, Universidade Estadual Paulista, Marília
e-mail: placida@marilia.unesp.br

Resumo

A relevância tanto da *web* quanto das bibliotecas digitais para os diversos ramos da ciência tem impulsionado pesquisadores e comunidades científicas a buscar soluções de integração, intercâmbio e entendimento semântico sobre os conteúdos que nelas circulam. Neste contexto, a proposição da pesquisa é verificar, na literatura científica, a aplicabilidade e a funcionalidade das tecnologias e dos instrumentos disponíveis como padrões de representação e descrição de conteúdo e forma, no âmbito das bibliotecas digitais. O objetivo é identificar e caracterizar o uso e a aplicação do *MarcOnt* como ferramenta para construção de formas de representação da informação para bibliotecas digitais, na atualidade, destacando-o como instrumento para descrição dos aspectos semânticos dos conteúdos digitais nestas ambiências. Como metodologia, caracteriza-se por ser uma pesquisa de análise exploratória e descritiva, com o qual é possível obter subsídios para maior e mais detalhado entendimento e compreensão do *MarcOnt* e das tecnologias abarcadas no âmbito das bibliotecas digitais para construção da *web* semântica. Como resultado, procurou-se demonstrar a aplicação da ferramenta *MarcOnt* num ambiente informacional específico, a Biblioteca Digital Semântica *JeromeDL*, para analisar, nos fundamentos da catalogação descritiva, as formas de representação dos recursos informacionais. Assim, o padrão de descrição bibliográfica semântica *MarcOnt* oferece a possibilidade de melhoria na *performance* de interoperabilidade, não somente entre padrões de metadados, mas também entre sistemas e ambientes informacionais digitais.

Palavras-chave

MarcOnt. MARC21. *Web* semântica. Bibliotecas digitais. Representação da informação. Catalogação automatizada. Padrões de metadados. Informação e tecnologia.

Abstract

The relevancy of not only the web but also the digital libraries for the diverse branches of science has stimulated researchers and scientific communities to search for integration solutions, interchange and semantic understanding about the contents that circulate in them. In this context, the proposal of this research is to verify, in the scientific literature, the applicability and functionality of the technologies and instruments available as patterns of representation and description of content and form, in the ambit of digital libraries, with the objective of

¹ Dissertação premiada com o segundo lugar, Enancib, João Pessoa, ano 2009, sob a orientação da Professora Doutora Plácida Leopoldina Ventura Amorim da Costa Santos.
Pesq. bras. ci. inf., Brasília, v.3, n.1,p.155-168, jan./dez. 2010

identifying and characterizing the use and the application of MarcOnt as a tool for the construction of forms of information representation for digital libraries at the present time, highlighting it as an instrument for describing the semantic aspects of the digital contents in these environments. As a methodology, the present work is characterized by being an exploratory and descriptive analysis of the subject, with which it is possible to obtain subsidies for a better and more detailed understanding and comprehension of MarcOnt and the technologies that are included in the ambit of digital libraries for the semantic web construction. As a result, it was attempted to demonstrate the application of the MarcOnt tool in a specific informational environment, the JeromeDL Semantic Digital Library, to analyze the fundamentals of descriptive cataloguing the forms of representation of informational resources. Therefore, the MarcOnt semantic bibliographic description pattern offers the possibility of assuring an improvement in the interoperability performance, not only among metadata patterns, but also among digital informational systems and environments.

Key words

MarcOnt. MARC21. Semantic web. Digital libraries. Representation of information. Automatized cataloguing. Metadata patterns. Information and technology.

INTRODUÇÃO

A nossa sociedade é marcada atualmente por constantes mudanças nos mais variados segmentos, uma delas de cunho tecnológico, no contexto que conhecemos hoje por sociedade da informação. Nos últimos tempos, houve um aumento desordenado e caótico na quantidade de informações produzidas e disponibilizadas em meio digital, requerendo uma mudança e um repensar nas formas de armazenamento, de representação, de descrição e de preservação dos recursos informacionais digitais.

A *web* atual, caracterizada como espaço aberto e de fácil socialização do conhecimento, vem enfrentando mudanças significativas em sua estrutura e na forma de representar e apresentar os recursos informacionais digitais. Isto exige e pressupõe um novo olhar de diversos profissionais, especialmente, os da área de Ciência da Informação (CI) e de Biblioteconomia, que mantêm um papel *sine qua non*, como agentes transformadores, no século XXI.

A relevância tanto da *web* quanto das bibliotecas digitais para os diversos ramos da ciência tem impulsionado pesquisadores e comunidades científicas a buscarem soluções de integração, intercâmbio e entendimento semântico sobre os conteúdos que nelas circulam, a fim de proporcionar recuperação mais precisa, relevante e significativa para o usuário final.

No cenário atual, é possível verificar forte tendência para disponibilização de conteúdos digitais e falta de uso de padrões adequados, a fim de representar e descrever tais recursos em ambientes digitais, bem como garantir sua apresentação para o uso, a preservação e o (re)uso das informações para os usuários.

Na tentativa de encontrar caminhos para a solução do problema, Kruk e Decker e Zieborak (2005) apontam a *MarcOnt Initiative*. Trata-se de recurso que apresenta instrumentos subjacentes à *web* semântica, tais como ontologias e uso de padrões de metadados, contemplando parâmetros que vão ao encontro dos objetivos das bibliotecas digitais e propiciam meios mais adequados de representar e organizar os recursos informacionais digitais atrelados aos aspectos semânticos dos conteúdos.

Neste contexto, a pesquisa objetiva identificar ferramentas para construção de formas de representação de recursos em bibliotecas digitais e na *web*, de modo geral, com o intuito de caracterizar a *MarcOnt Initiative* como padrão de representação da informação (forma e conteúdo) em ambientes digitais. Pretende, também, justificar a adoção da ferramenta *MarcOnt* como possibilidade de proporcionar a interoperabilidade semântica entre bibliotecas digitais e outros sistemas de informação.

Por meio de metodologia exploratória e descritiva do tema, identificou-se o embasamento conceitual subjacente às ferramentas para a padronização da descrição dos aspectos de forma e conteúdo dos recursos informacionais nestas ambiências, na abordagem da CI.

PANORAMA DA CATALOGAÇÃO DESCRITIVA EM TEMPOS DIGITAIS

Pode-se observar, contemporaneamente, um cenário marcado por grandes transformações tecnológicas e pelo surgimento e criação das mais variadas técnicas para as questões de representação, descrição, produção, organização, transmissão e uso das informações em ambientes informacionais digitais.

No entanto, somente parte dessas informações e em ambientes específicos, tais como bibliotecas digitais, por exemplo, possui um padrão de metadados para a descrição bibliográfica que possibilita a padronização dos recursos informacionais para atender aos requisitos da interoperabilidade.

Neste contexto tecnológico, é possível observar o surgimento de novas tecnologias baseadas em pesquisas na temática *web* semântica, com o intuito de tentar agregar semântica à descrição bibliográfica para tratamento mais efetivo dos conteúdos digitais. A iniciativa *MarcOnt* surge, então, na tentativa de solucionar os problemas de descrição e de representação dos recursos informacionais em ambientes digitais, tais como bibliotecas digitais, atrelada aos padrões de metadados advindos dos campos da Biblioteconomia e da CI. Bom exemplo é o *Machine Readable Cataloging* (MARC21) para a descrição bibliográfica e outros padrões, como o *Dublin Core*, para a localização de recursos na *web*.

A *MarcOnt Initiative* se desenvolveu a partir da Biblioteca Digital Semântica *JeromeDL*. Num primeiro momento, as primeiras experiências na *JeromeDL* eram positivas. Logo depois, foram detectados alguns problemas, pois os objetivos iniciais eram desenvolver e incrementar a busca automaticamente, atribuindo semântica aos recursos informacionais a fim de potencializar os processos de recuperação de informação. (SYNAK, 2005). Diante dos entraves surgidos, esforços colaborativos foram empreendidos pelos pesquisadores da *JeromeDL* e pela equipe envolvida no projeto. Decidiu-se, à época, tratar esses problemas criando-se uma ontologia bibliográfica que contemplasse os aspectos semânticos de forma e de conteúdo dos recursos informacionais e que atuasse, de forma mais efetiva, numa biblioteca digital semântica. Assim, os requisitos iniciais foram elaborados por Sebastian Ryszard Kruk e se constituiu o projeto *MarcOnt Initiative*©2005, o qual pode ser encontrado, de forma detalhada, no Portal *MarcOnt*, <http://portal.marcont.org>.

Segundo Kruk e Synak e Zimmermann (2005a, p. 1, tradução nossa), o objetivo central do *MarcOnt* é “[...] desenvolver um novo padrão de descrição bibliográfica sob a forma de uma ontologia e ferramentas relacionadas da *web* semântica utilizando tecnologias semânticas”. Castro (2008, p. 124) complementa e afirma que a *MarcOnt Initiative* é:

[...] um padrão de descrição bibliográfica semântica que contempla forma e conteúdo do recurso informacional, a fim de sua identificação e individualização para leitura e processamento compreensíveis por máquina / computador, utilizando-se de instrumentos e ferramentas tecnológicas subjacentes à *web* semântica, podendo ser aplicado no contexto de ambientes informacionais digitais.

O primeiro passo para agregar semântica à descrição das informações em bibliotecas digitais e sua compatibilidade para que ocorra a interoperabilidade é o desenvolvimento de uma ontologia comum. A ontologia pode ser considerada um instrumento para a descrição bibliográfica dos recursos informacionais, pois se cria uma rede de conceitos com propriedades apropriadas e restritas. Kruk e Synak e Zimmerman (2005b, p. 1, tradução nossa) alegam que “[...] infelizmente, é muito difícil convencer os bibliotecários a realizarem o esforço da criação da descrição semântica dos recursos, por exemplo, os livros”.

Aqui, destaca-se a relevância da participação ativa de profissionais de áreas interdisciplinares na criação da iniciativa *MarcOnt*, em especial, dos profissionais da Biblioteconomia, uma vez que este profissional, por formação, domina técnicas e habilidades necessárias ao tratamento documentário dos recursos informacionais. A catalogação, instrumento de descrição bibliográfica, pautada em regras e esquemas de descrição, fornece subsídios para construção de formas de representação e descrição de forma padronizada, possibilitando armazenamento, preservação, uso e (re)uso das informações de forma mais efetiva.

Como Synak (2005, p. 50, tradução nossa) aponta, muitas são as razões para construção de um novo padrão para a descrição bibliográfica semântica. Dentre elas, pode-se citar o argumento de que o padrão de metadados MARC21 não fornece informação semântica necessária de forma a ser compreendida por uma máquina:

O formato MARC21 nos permite descrever dois tipos específicos de recursos: Língua dos materiais (que pode ser aproximadamente descrito como livros, periódicos, jornais etc.) e Língua de materiais manuscritos (tais como livros valiosos, livros antigos, escritos à mão).

Em outras palavras, quando o autor se refere à Língua dos materiais e Língua de materiais manuscritos, é possível se reportar à estrutura do MARC na posição **06** (tipo de registro) do LÍDER, onde se utiliza a letra “a” para material textual (impresso) e a letra “t” para material textual (manuscrito).

Um recurso que represente a “Língua dos materiais” é identificado em MARC21 colocando a letra “a” no campo número 06 no líder do registro MARC21. Para a “Língua de materiais manuscritos”, a marca apropriada requer colocar a letra “t” no mesmo lugar. O problema é que não há absolutamente nenhuma conexão entre a “Língua do material” e a “Língua de material manuscrito”. Para alguém que não tem uma pista quando se trata de livros, estes dois tipos de recursos poderiam ser chamados “Pedras” e “Flores”. Este é o caso quando uma máquina é usada para processar o dado. (SYNAK, 2005, p. 51, tradução nossa).

Neste mesmo sentido, Synak (2005, p. 51, tradução nossa), para explicação ainda mais plausível, afirma que:

MARC21 não nos permite descrever que a “Língua dos materiais” e a “Língua de materiais manuscritos” são ambas adequadas para leitura (têm algumas características comuns) ou que a “Língua de materiais manuscritos” é de fato também, a “Língua dos materiais”, mas têm algumas características específicas que as fazem diferentes. Se quiséssemos escrever um termo numa máquina de busca que incluísse nessa busca manuscritos na falta de livros usuais sobre o assunto, teríamos que especificar tal conexão. Uma ontologia nos permite fazer isso automaticamente.

Na busca por subsídios, atrelada às motivações da execução desta pesquisa que procuramos responder e caracterizar, o *MarcOnt* está pautado na tese que o autor procura defender. Argumenta que, embora o MARC21 seja um padrão mundialmente e internacionalmente aceito e adotado pelas bibliotecas e pela comunidade da Biblioteconomia, de modo geral, possuindo estrutura completa e complexa, não há um padrão de descrição semântica que contemple a profundidade da descrição, ou que cubra a extensão requerida ou suportada pelo MARC21 usado para os propósitos da Biblioteconomia.

Pensando assim, pode-se considerar que o padrão de descrição bibliográfica semântica *MarcOnt* pode possibilitar melhoria na explicitação das informações, de forma a permitir a compreensão, pela máquina, das solicitações de busca para o usuário final. Synak (2005) questiona, também, a necessidade da razão ao se criar outra ontologia, dizendo que ainda não existe uma adequada que vá ao encontro dos objetivos da descrição bibliográfica e que atenda aos requisitos dos profissionais da Biblioteconomia.

As novas soluções não podem negligenciar padrões e formatos existentes ou nunca se tornarão amplamente conhecidas. No momento, o formato de dados bibliográficos MARC21 é ainda o principal padrão utilizado em bibliotecas de todo o mundo. Uma ontologia que aspira a ser seu sucessor (ou ao menos seu equivalente semântico) não pode oferecer menor capacidade na descrição. (SYNAK, 2005, p. 63, tradução nossa).

Na tentativa de responder a alguns questionamentos sobre a criação de uma ontologia bibliográfica, bem como sua relevância, o autor menciona que se vivenciou uma experiência com descrições semânticas na Biblioteca Digital Semântica *JeromeDL* e que os resultados apontam incremento nos serviços de busca e de comunicação entre bibliotecas digitais, favorecendo amplitude na recuperação dos recursos informacionais. Vale destacar que, para o desenvolvimento pleno do *MarcOnt*, os estudos mostram e apontam exigências ou requisitos para atingir suas funcionalidades. Por isto, seu idealizador procurou, por meio do estado-da-arte no domínio da Biblioteconomia, levantar os principais padrões de metadados na área, visando descobrir se haveria compatibilidade com a iniciativa *MarcOnt*. Segundo Synak e Kruk (2005), dentre os principais requisitos, destacam-se:

- ◆ A ontologia deve prover não menos capacidade de descrição do que o formato de descrição bibliográfica MARC21.
- ◆ A ontologia deve ser descrita usando *Web Ontology Language* (OWL DL).
- ◆ A ontologia deve reusar (se possível) outras ontologias existentes ou suas partes. Deve ser relativamente fácil para aderir a ontologia *MarcOnt* a outras existentes, se requerido.
- ◆ Deve ser fácil traduzir descrições semânticas *MarcOnt* compatíveis com outros formatos, tais como MARC21, Dublin Core ou BibTeX.

- ◆ O primeiro esboço / esquema da ontologia deve ser ponto de partida para trabalhos futuros.

Destaca-se que o *MarcOnt* já está disponível na Biblioteca Digital Semântica *JeromeDL*, o que será tratado nas seções seguintes. O próprio Portal *MarcOnt* disponibiliza um ambiente para construção de descrições bibliográficas semânticas para experiências e testes.

INTEROPERABILIDADE ENTRE PADRÕES DE METADADOS: A ARQUITETURA *MARCONT*

Kruk e Synak e Zimmermann (2005a, 2005b), para melhor compreensão do padrão de descrição semântica *MarcOnt*, propuseram uma arquitetura inicial, a fim de fornecer uma visualização de como seria seu funcionamento ao se pensar na questão da interoperabilidade semântica entre padrões de metadados heterogêneos, bem como o acesso aos recursos informacionais numa única interface.

A este respeito, a **Figura 1** apresenta o delineamento de uma arquitetura inicial dos serviços de mediação *MarcOnt*.

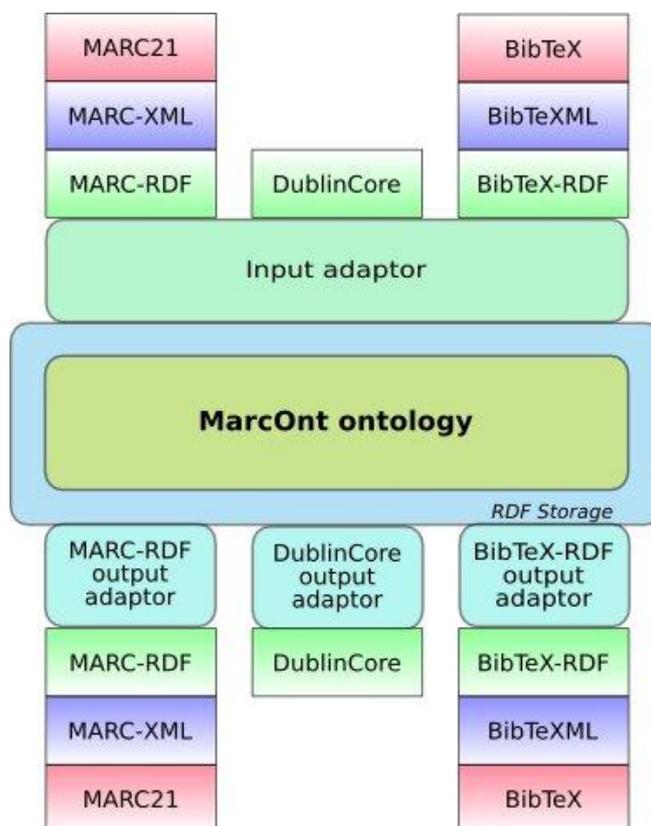


FIGURA 1 – Arquitetura de serviços de mediação *MarcOnt*
Fonte: Kruk, Synak e Zimmermann (2005a, p. 2)

A arquitetura forma um fluxo, onde a parte superior é a entrada (*input*), as descrições semânticas estão no centro (*MarcOnt Ontology*) e a parte mais inferior é a saída (*output*). Nota-se que o centro da arquitetura é o *Resource Description Framework* (RDF), considerado o núcleo de armazenamento para descrições semânticas compiladas de outros padrões de metadados.

As caixas na primeira fileira (na entrada) e na última fileira (na saída) da arquitetura representam os formatos inerentes aos tradicionais (arquivos neste formato). As caixas no meio (MARCXML, BibTeXML) representam a primeira fase do processo de tradução, onde os arquivos inerentes são analisados gramaticalmente (sintaxe) e seus conteúdos são incluídos no formato XML.

Os adaptadores de entrada e saída (*input* e *output*) são implementados usando o mecanismo de inferência Sesame. Trata-se de uma ferramenta (mais informações em <<http://www.openrdf.org>>), ou seja, um conjunto de classes-base desenvolvido em Java com código aberto para o desenvolvimento de sistemas, com o fim de armazenar e consultar dados em RDF. Este sistema é totalmente extensível e configurável no que concerne aos mecanismos de inferência e ao armazenamento de dados em formato RDF. Retomando o mecanismo de inferência Sesame, acrescenta-se que, de um lado dos adaptadores, estão as descrições semânticas e do outro, dados de descrições bibliográficas convertidas para um modelo RDF. Converter diferentes formatos de descrição para o modelo RDF depende do formato original do arquivo. Registros em formato MARC21 são armazenados em arquivos binários. Portanto, a conversão exige primeiro que tais arquivos sejam traduzidos para o formato MARCXML e, numa fase posterior, convertem-se os dados para o modelo RDF usando XSLT.

Vale ressaltar que o autor destaca que ainda não há representações oficiais de dados de MARC21 em RDF, desenvolvido pela *Library of Congress* (LC), e que os pesquisadores do *MarcOnt* criaram seu próprio formato (MARCRDF). Logo, nas palavras de Synak (2005, p. 75, tradução nossa), “apoiar outros formatos é uma questão de desenvolver adaptadores e conversores apropriados. Atualmente, estamos desenvolvendo adaptadores para MARC21”.

Ainda segundo o autor supra, a fase final requer traduzir os dados RDF para estruturas compatíveis com o padrão de descrição bibliográfica *MarcOnt*. Os estágios precedentes demandam escrever ferramentas simples para que as bibliotecas as usem e possam intercambiá-las com relativa facilidade. Assim, pode-se dizer que essa arquitetura inicial visa ao compartilhamento entre padrões de metadados distintos, e pode proporcionar interoperabilidade semântica entre bibliotecas digitais, por exemplo, onde as informações semânticas estarão disponíveis e armazenadas numa única interface, neste caso, no *MarcOnt*.

Outra releitura sobre a questão da descrição bibliográfica, pautada na arquitetura de serviços de mediação *MarcOnt*, pode ser compreendida de acordo com a **Figura 2**. Reforçam-se a criação e a necessidade de um padrão de descrição bibliográfica semântica que contemple a representação de forma e de conteúdo dos recursos informacionais.

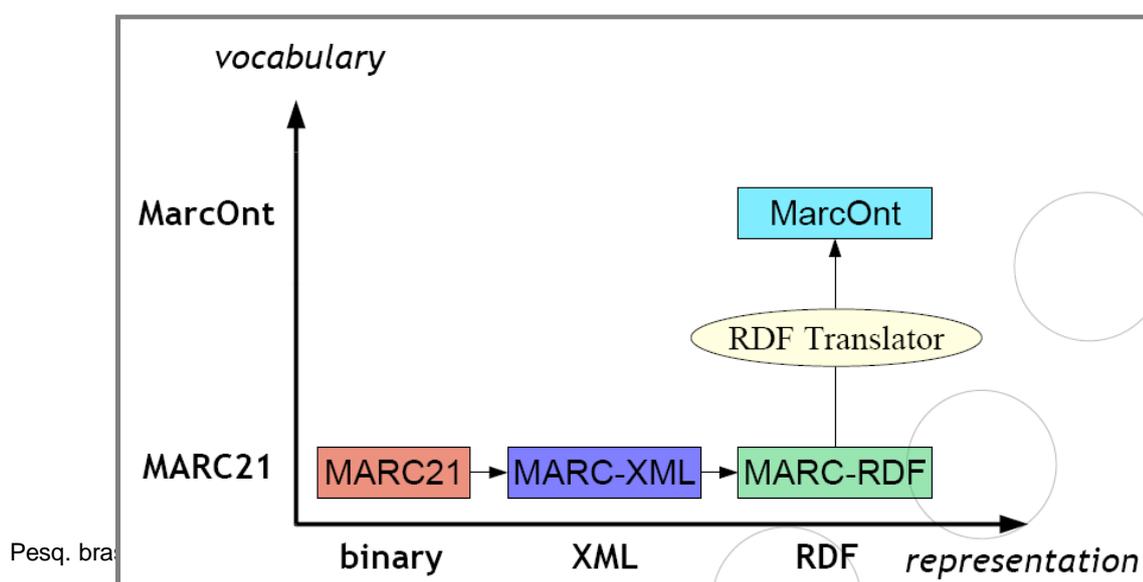


FIGURA 2 – Estrutura de representação em *MarcOnt*
 Fonte: Kruk, Synak e Zimmermann (2005a, p. 27)

Conforme se observa na **Figura 2**, há representação dos recursos informacionais no padrão de metadados MARC21, MARCXML (na estrutura MARC com a linguagem computacional XML) e MARC-RDF (estrutura MARC em RDF), uma vez que essas representações passam por ferramentas de tradução / conversão, dadas por *RDF Translator*, gerando informações semânticas armazenadas em *MarcOnt*. Assim, pode-se dizer que o estabelecimento de uma linguagem comum (*vocabulary*), que proporcione e garanta busca e recuperação de recursos informacionais para uso e reuso das informações pelo usuário final será determinado por *MarcOnt*.

Um das grandes dificuldades apontadas pela literatura e um problema para a interoperabilidade entre padrões de metadados distintos em bibliotecas digitais é o fato de que os dados para a tradução das informações nem sempre estão representados em RDF. Isto implica o desenvolvimento de novas tecnologias para o tratamento adequado dos recursos informacionais.

Além do mais, Synak (2005, p. 75, tradução nossa) menciona a dificuldade em se criarem regras que contemplem os objetivos do *MarcOnt*: “[...] a tarefa requer escrever centenas de regras de inferência para traduzir aos usuários do *MarcOnt* sua lógica de funcionamento, bem como suas propriedades para garantir de forma apropriada um pouco mais de informação sem possíveis perdas”. Em sua visão, os formatos de descrição bibliográfica são exatamente tal caso. A informação é armazenada em cada arquivo binário ou textual, a depender do formato. Para processar a informação usando as tecnologias da *web* semântica, é necessário que, primeiramente, representemos estas informações em RDF.

Ainda neste sentido, Synak (2005) enumera duas fases para a tradução / conversão entre os padrões de metadados:

1. Tradução do formato original para RDF.
2. Tradução de um grafo RDF para outro.

Os três formatos avaliados são MARC21, BibTeX e *Dublin Core*.

1. MARC21 – formato binário. Os arquivos necessitam ser analisados gramaticalmente. O resultado de um arquivo no formato MARCXML pode ser transformado em RDF usando *XSL Transform*. O problema é que não há nenhuma representação oficial de MARC em RDF, ao menos, não desenvolvido pela LC. Para nossas necessidades, especifica-se o formato MARCRDF (para *RDF Schema* de MARCRDF), representação RDF das estruturas de MARCXML.
2. BibTeX – formato de texto. Segue a sintaxe TeX. Os arquivos BibTeX necessitam ser analisados gramaticalmente antes de ser processados mais adiante. Há representações BibTeX em XML (BibTeXXML) e RDF (BibTeX em OWL). O arquivo XML

analisado, gramaticalmente, pode ser convertido, com facilidade, para RDF usando XSLT.

3. Dublin Core – muitas sintaxes disponíveis (em XML, como metadado em HTML, RDF). Os elementos de DC são usados geralmente em RDF. Ao contrário do MARC21 ou BibTeX, DC não é um formato de arquivo fechado, podendo ser mais flexível na transformação dos dados. DC é mais propriamente um conjunto de elementos que podem ser usados onde forem aplicáveis. (SYNAK, 2005, p. 73, tradução nossa).

Nessa perspectiva, foram formuladas algumas regras para transformações de descrições no padrão de metadados MARC21, a fim de possibilitar melhores níveis de relacionamento entre os registros bibliográficos. A **Figura 3** apresenta regras que poderiam ser usadas no processo de transformações de descrições em MARC21 para descrições semânticas em *MarcOnt*. No exemplo, constam regras para um registro MARC21, descrevendo um documento do tipo livro, com metadados de autor e título, utilizando os conceitos de *MarcOnt* (*Marcont:Book*) para o armazenamento das informações.

<i>Rule ID</i>	<i>Premises</i>	<i>Consequents</i>
1	LDR/06 has value of "a"	create individual of marcont:Book call rule 2 call rule 4
2	datafield 245 exists	create property marcont:hasTitle pointing to: create individual of marcont:TitleStatement call rule 3
3	subfield 245\$a exists	create property marcont:titleValue with value from 245\$a
4	datafield 100 exists	create property marcont:hasCreator pointing to: create individual of foaf:Person call rule 5
5	subfield 100\$a exists	create property foaf:Surname with value from 210\$a
6

FIGURA 3 – Regras para construção de descrição semântica de MARC21 para *MarcOnt*
Fonte: Synak (2005, p. 74)

Vale lembrar que o padrão de metadados MARC21 descreve vários recursos informacionais e que o exemplo da **Figura 3** (livros) é apenas um dos tipos documentais contemplados pelo formato. Verifica-se, ainda, a criação de regras para apenas dois campos específicos abarcados pelo padrão de metadados MARC21 (autor e título), o que implica, para um registro completo, a construção detalhada e padronizada de normas que forneçam a semântica necessária para o entendimento de cada campo para leitura das informações pela máquina / computador.

A ontologia terá de ser atualizada constantemente para permitir melhor representação da informação. As regras existentes terão que refletir tais mudanças. Atualmente, existe um conjunto de regras que permite mover uma quantidade limitada de informações de formatos tradicionais para *MarcOnt*. A informação mais importante, como título ou autor, é suportada. Criar ferramentas que simplificariam e acelerariam o processo de criação de regras será a próxima etapa e um dos assuntos de trabalhos no Portal Colaborativo *MarcOnt*. A primeira

etapa realizada foi criar um *framework* (estrutura), uma máquina (sistema), que poderia processar as regras. (SYNAK, 2005, p. 74, tradução nossa).

Assim, tratar-se-á, agora, da aplicação do *MarcOnt* num ambiente informacional específico, a Biblioteca Digital Semântica denominada *JeromeDL*, onde se pode verificar sua aplicabilidade e funcionalidade, fornecendo modelos de experiências e de testes.

JEROMEDL – MODELO DE AMBIENTE COLABORATIVO DE INTELIGÊNCIA COLETIVA

Para acompanhar o cenário de evoluções tecnológicas e como modelo de aplicação do padrão de descrição semântica *MarcOnt*, surge a Biblioteca Digital Semântica *JeromeDL* como ambiência que busca, por meio da semântica, possibilitar refinamento nos mecanismos de busca. Originou-se de experiências e testes da tese de Sebastian Ryzard Kruk na *Gdańsk University of Technology (GUT)*, Polônia, anos 2004 / 2005, contando com a participação e a supervisão da *Digital Enterprise Research Institute (DERI)*, em Galway, na Irlanda, segmento da *National University of Ireland*, na mesma cidade. Enfim, *JeromeDL* consiste em biblioteca digital semântica que utiliza as principais tecnologias abordadas na *web* semântica.

A descrição dos recursos informacionais tem como base a arquitetura de metadados RDF e a realização de buscas semânticas baseadas em ontologias, o que significa a possibilidade de melhorias na busca e na usabilidade, com o estabelecimento de maior nível de interoperabilidade entre ambientes informacionais digitais. (KRUK; DECKER; ZIEBORAK, 2005, tradução nossa).

A Biblioteca Digital Semântica *JeromeDL* foi projetada a fim de atender desde os profissionais responsáveis pela confecção dos recursos informacionais até os usuários que irão se beneficiar e fazer uso desses recursos. Assim, de acordo com Kruk e Decker e Zieborak (2005, p. 2, tradução nossa), foram traçados os principais requisitos para *JeromeDL*, tais como:

- ◆ Compreender a herança de bibliotecas clássicas / convencionais (exemplo: livros antigos).
- ◆ Prover características de navegação orientadas ao usuário.
- ◆ Permitir busca eficiente.
- ◆ Abarcar múltiplos formatos de recursos.
- ◆ Possibilitar a comunicação com outros sistemas de biblioteca digital.
- ◆ Utilizar resultados das últimas pesquisas na *web* semântica e o gerenciamento da comunicação e da informação nas pesquisas.

Ao se pensar em formas de representação da informação registrada, além da descrição e da representação dos recursos num ambiente informacional, de acordo com normas e padrões que norteiam sua construção, deve-se, também, pensar nas formas de apresentação

do recurso informacional, de forma a possibilitar boa usabilidade dos sujeitos que se apropriarão das informações no sistema.

A **Figura 4** mostra uma interface da Biblioteca Digital Semântica *JeromeDL*, apresentando os recursos informacionais no momento de busca e sua visualização num padrão de metadado de acordo com a preferência do usuário. Verifica-se um recurso informacional apresentado ao usuário numa única plataforma contemplando os padrões de metadados vislumbrados na iniciativa *MarcOnt* (BibTeX, BibTeXML, BibTeXRDF, Dublin Core, MARC21, MARCXML, MARCRDF e *MarcOnt*), uma vez que há, nessa interface, a possibilidade de interoperabilidade entre padrões de metadados heterogêneos.

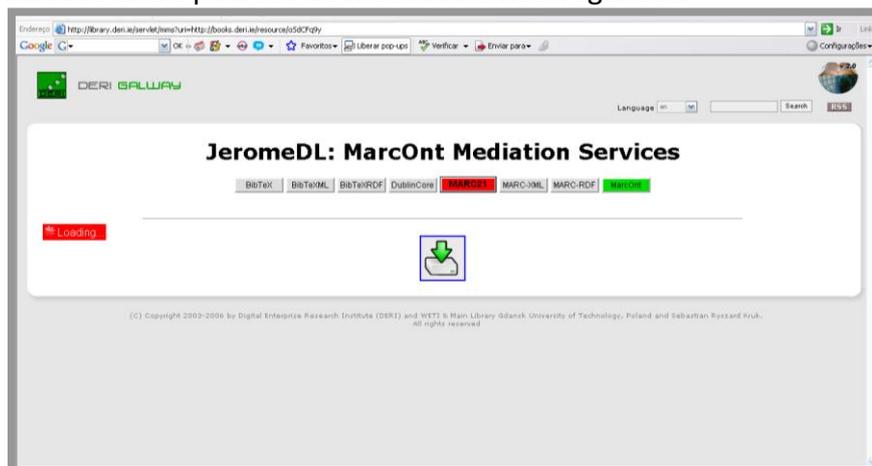


FIGURA 4 – Formas de apresentação dos recursos informacionais em diferentes tipos de padrões
 Fonte: Disponível em: <<http://library.deri.ie/servlet/mms?uri=http://books.deri.ie/resource/oSdCFq9y>>. Acesso em: 31 out. 2007.

BREVES PONTUAÇÕES ACERCA DA REPRESENTAÇÃO E DA DESCRIÇÃO DOS RECURSOS INFORMACIONAIS NA *JEROMEDL* À LUZ DA CATALOGAÇÃO DESCRITIVA

Faz-se necessário tecer alguns comentários, pautados e vivenciados na prática biblioteconômica, mais especificamente à luz dos fundamentos da catalogação, que permite a padronização da descrição e da representação dos recursos informacionais.

Um dos pontos que vale a pena destacar é como se apresenta uma descrição em MARC21, na Biblioteca Digital Semântica *JeromeDL*. Em outras buscas realizadas na Biblioteca, com a finalidade de verificar a representação e a descrição de um recurso informacional, foram observadas falhas nesses quesitos. Encontramos um registro MARC21 que possuía mais de um autor responsável pela obra. E o mais curioso era que o campo 100 (entrada principal – nome pessoal) do registro estava repetido. De acordo com a estrutura MARC, este campo não se repete, pois de acordo com os esquemas de descrição AACR2, que antevêm ou norteiam a construção do campo numa estrutura legível por máquina MARC, seria necessária a criação de uma entrada secundária de autoria, que no registro MARC é dada pelo campo 700 (entrada secundária – nome pessoal). Vale lembrar, também, que até o presente momento, não se encontram disponíveis representação e descrição em MARC21 em sua totalidade, sendo relevantes, segundo Kruk (2006) apenas os campos de autoria (100) e título (245) para busca e recuperação de recursos informacionais. Desta forma, o bibliotecário / catalogador é considerado protagonista na construção de regras para o estabelecimento da semântica para outros campos do registro MARC21, podendo estes compartilhar informações semânticas não só num mesmo registro / estrutura, mas, também, com outros padrões de metadados e em sistemas diferentes.

Desta forma, reforça-se que, para estabelecimento mais efetivo de representações na estrutura *MarcOnt* e para a funcionalidade da Biblioteca Digital Semântica *JeromeDL*, é imprescindível a participação ativa e direta do profissional bibliotecário. Como dito, é ele quem detém técnicas e habilidades necessárias ao tratamento das informações de forma padronizada e pautada em normas e regras na construção e confecção de registros bibliográficos. Em consonância com profissionais da Computação e com outros atores envolvidos no processo, o bibliotecário pode propiciar a clareza nos serviços oferecidos pela biblioteca digital semântica, no atendimento e na satisfação nas solicitações de buscas requeridas pelo usuário.

À GUIA DE CONSIDERAÇÕES E REFLEXÕES

Atualmente, vivencia-se uma mudança de paradigmas, e novos horizontes se abrem à catalogação descritiva, bem como novas tendências e novos olhares para os registros bibliográficos no século XXI.

O padrão de descrição bibliográfica semântica *MarcOnt* faz vislumbrar nova tendência para os ambientes informacionais digitais de modo a permitir a descrição semântica dos recursos informacionais e a possibilidade de interoperabilidade semântica, não somente entre os campos de um registro bibliográfico, mas, ainda, entre outros padrões de metadados em plataformas e sistemas heterogêneos.

A relação entre outros padrões de metadados, como MARC21 e *MarcOnt*, não tem como meta transformar a estrutura preestabelecida do MARC21 na linguagem da ontologia, mas fornecer a informação semântica à descrição dos recursos, de modo a oportunizar compreensão pela máquina / computador.

O padrão de descrição semântica *MarcOnt* é um instrumento tecnológico que, no contexto das bibliotecas digitais e da *web* de modo geral, pretende atuar como ferramenta capaz de oportunizar ao tratamento dado aos recursos informacionais descrição de forma e de conteúdo legível por máquinas. Isto trará a possibilidade de resultados compreensíveis aos humanos para apresentação das informações com vistas à recuperação mais eficiente.

Há, no padrão de descrição bibliográfica semântica *MarcOnt*, um ponto de partida para integração de fontes / sistemas que se utilizam de padrões de metadados heterogêneos, no que tange à melhoria da *performance* na manipulação de formas de representação bibliográfica e da interação dos sujeitos psicossociais com o ambiente informacional. Para tanto, pode-se considerar que tal padrão de descrição bibliográfica semântica:

- ◆ Apresenta-se como novo padrão para a representação (forma e conteúdo) do recurso informacional no contexto digital, na busca por uma descrição bibliográfica mais completa e complexa, com possibilidade de garantir aperfeiçoamento na *performance* de interoperabilidade entre sistemas e ambientes informacionais digitais.
- ◆ Pode possibilitar melhora na explicitação das informações de forma a permitir compreensão pela máquina nas solicitações de busca para o usuário final.
- ◆ Pode possibilitar compartilhamento de recursos entre ambientes informacionais heterogêneos, indo ao encontro da filosofia colaborativa da *web* semântica e, recentemente, *web 2.0*.

- ◆ Apresenta-se como ferramenta tecnológica que pode potencializar a padronização das informações na otimização dos processos de representação, descrição, organização, acesso, recuperação, preservação, uso e (re)uso dos recursos informacionais em distintas ambiências.

Assim, pode-se dizer que, para o funcionamento pleno do padrão de descrição semântica *MarcOnt*, é necessário que este compreenda e absorva a égide dos conceitos oriundos da catalogação descritiva, uma vez que esta proporcionará e subsidiará elementos fundamentais para construção padronizada da representação. Além disto, existe a possibilidade de permitir avanço na *performance* de interoperabilidade semântica entre os padrões de metadados vislumbrados em *JeromeDL* e outros sistemas e ambientes informacionais digitais, de forma a garantir a apresentação dos recursos para preservação, uso e (re)uso das informações para o usuário final.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pelo suporte financeiro na elaboração da pesquisa.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) pela continuação em pesquisa de doutoramento.

REFERÊNCIAS

CASTRO, F. F. de. **Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da Ciência da Informação**: uma abordagem do MarcOnt Initiative na era da web semântica. 2008. 201 f. Dissertação (Programa de Pós-Graduação em Ciência da Informação) – Faculdade de Filosofia e Ciências da Universidade Estadual Paulista, Marília. 2008.

KRUK, S. R. **JeromeDL 2.0**: a digital library on social semantic information spaces. 2006. Disponível em: <http://www.sebastiankruk.com/storage/presentation/jeromedl2.Ostable_gut/img0.html>. Acesso em: 25 ago. 2007.

KRUK, S. R.; DECKER, S.; ZIEBORAK, L. **JeromeDL**: reconnecting digital libraries and the semantic web. 2005. Disponível em: <http://www.marcont.org/marcont/pdf/www2005_jeromedl.pdf>. Acesso em: 14 jan. 2006.

KRUK, S. R.; SYNAK, M.; ZIMMERMANN, K. **MarcOnt Initiative**: mediation services for digital libraries. 2005a. Disponível em: <<http://www.marcont.org/marcont/pdf/marcontecd12005.pdf>>. Acesso em: 16 jan. 2006.

KRUK, S. R.; SYNAK, M.; ZIMMERMANN, K. **MarcOnt Initiative**: integration ontology for bibliographic description formats. 2005b. Disponível em: <<http://dc2005.uc3m.es/program/presentations/Thursday%2015.%2015.30%20h%20-%20s.kruk.pdf>>. Acesso em: 16 jan. 2006.

SYNAK, M. **MarcOnt Ontology**: semantic MARC21 description for L2L & L2C communication. 2005. 126 f. Thesis (Master in Informatics, Distributed Computer Systems) – Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk. 2005.

SYNAK, M.; KRUK, S. R. **MarcOnt Initiative**: the ontology for the librarian world. 2005. Disponível em: <http://www.marcont.org/marcont/pdf/ms_eswc2005marcont.pdf>. Acesso em: 6 jan. 2008.