



## Scan for Marc: conversão de registros em fichas para o Formato MARC21 Bibliográfico<sup>1</sup>

Zaira Regina Zafalon

Plácida Leopoldina Ventura Amorim da Costa Santos

Ana Maria Pereira, Jairo da Silva

Jairo da Silva

**Resumo:** Aborda conversão de registros bibliográficos a ser usado por instituições que não disponibilizam a totalidade de dados de seus acervos em catálogos online de acesso público. Apesar da possibilidade de recorrerem a processos de conversão retrospectiva de registros bibliográficos para saldar tal lacuna, trabalhos técnicos tornam-se necessários, tais como adequação de pontos de acesso de assunto, de notas e outras informações locais nos quais incorrem correção, supressão de campos, modificação de conteúdo e acréscimo de informações locais. Apresenta-se a conversão retrospectiva de dados bibliográficos a partir do processamento de imagens, que considera o aproveitamento integral do registro bibliográfico da própria instituição (dados bibliográficos, pontos de acesso e dados de localização, já definidos nos registros analógicos). Propõe-se como objetivo principal apresentar o Scan for MARC, um interpretador computacional de acesso web para conversão de dados bibliográficos analógicos para o Formato MARC21 Bibliográfico. Recorre-se à abordagem qualitativa, com caráter exploratório e descritivo, e à pesquisa bibliográfica. Os resultados apresentam contribuições no plano científico, diante do desenvolvimento de aspectos sintáticos e semânticos de registros bibliográficos; no plano tecnológico, dado o desenvolvimento do Scan for MARC como produto para a melhoria de processos de conversão de dados bibliográficos analógicos; no plano social, inerente ao compartilhamento de registros e às suas metodologias, à ampliação do acesso aos metadados e à interoperabilidade entre sistemas gerenciadores de bibliotecas, fatores que contribuem para a discussão e socialização do conhecimento científico/tecnológico.

---

<sup>1</sup> A presente pesquisa é desenvolvida no âmbito do Grupo de Pesquisa Tecnologias em Ambientes Informacionais – GPTAI, certificado pela Universidade Federal de São Carlos (UFSCar) e pelo Conselho Nacional de Pesquisa Científica (CNPq). Este artigo apresenta-se como resultado parcial de projeto de pesquisa aprovado junto ao CNPq, sob número 472167/2014-6, e caracteriza-se como sendo continuidade de outra, iniciada em 2009, que culminou na conclusão da pesquisa de doutorado junto à Universidade Estadual Paulista Júlio de Mesquita Filho, no registro do Scan for MARC, junto ao INPI, ainda como interpretador computacional em linha de comando, e na publicação, pela Editora UNESP, do livro *Scan for MARC: conversão de registros bibliográficos analógicos para o Formato MARC21 Bibliográfico*, além de discussão em eventos.



**Palavras-chave:** Scan for MARC. Conversão retrospectiva. Catalogação; Automação de bibliotecas. Catálogos online de acesso público.

## INTRODUÇÃO

Esta pesquisa insere-se no contexto social contemporâneo, no qual se faz uso e aplicação de tecnologias no seu cotidiano. Aplicações tecnológicas em estoques informacionais apresentam vantagens como intercâmbio de dados, interoperabilidade entre sistemas e, por consequência, facilidade na descrição e no acesso aos dados bibliográficos.

As unidades de informação, em específico, nesta pesquisa, as bibliotecas, tem vivenciado a (re)construção do cenário de transformações socioculturais e tecnológicas. Como não poderia deixar de ser, tem se mostrado necessária a (re)avaliação de processos, produtos e serviços que, motivados pelas aplicações tecnológicas, tem influenciado e transformado a geração e o uso de informações. Tem sido alterados processos de descrição, tratamento, organização, distribuição, armazenamento e, em decorrência disso, a recuperação e o acesso às informações, tanto por humanos quanto por máquinas.

Com as novas tecnologias de informação e comunicação surgem, também, as facilidades de compartilhamento de registros bibliográficos entre as mais variadas unidades de informação.

O desafio, entretanto, está presente nas instituições que ainda não se utilizam deste universo de facilidades por conta de não fazerem uso de recursos tecnológicos, por quaisquer que sejam os motivos (e esta é uma especulação a partir de atividades profissionais, acadêmicas e extensionistas!): limitações institucionais, de infra-estrutura tecnológica; limitações pessoais, na compreensão e uso de tecnologias informáticas; inexperiência no processo de seleção de *software* de gerenciamento que atenda aos requisitos de compartilhamento de dados; inabilidade de gestão de diferenciados recursos necessários ao processo de mudança ou implantação de sistemas



automatizados; desconhecimento de políticas e ações necessárias para o compartilhamento, a conversão e a migração de dados bibliográficos, apesar de se ter conhecimento da existência de padrões e das vantagens na adoção. Por conta disso, concorda-se com a afirmação de Lancaster (1994, p. 8) de “que é a tecnologia disponível que determina como os serviços de biblioteca serão no futuro.”

O principal problema a ser abordado está presente no dilema de que ainda se encontram instituições que não conseguiram disponibilizar a totalidade de dados bibliográficos de seus acervos em catálogos online de acesso público, ou seja, os registros ainda compõem catálogos bibliográficos impressos em fichas (registros analógicos). Algumas bibliotecas nesta situação, e que se destacam por conta de seu acervo e de sua relevância social e de pesquisa, são: a Biblioteca Nacional, no Rio de Janeiro (RJ) e a Biblioteca Mario de Andrade, em São Paulo (SP).<sup>2</sup> A utilização de catálogos digitais, em detrimento dos catálogos analógicos, pode ser feita utilizando-se de processos de conversão retrospectiva de registros bibliográficos que, até os últimos anos, se configurou a partir ações de reaproveitamento de registros disponíveis em catálogos online em redes de cooperação.

É sobre este ponto que se apresenta a proposta desta pesquisa, uma vez que se indica uma nova forma de se proceder a conversão retrospectiva de dados bibliográficos que considera o aproveitamento integral do registro bibliográfico da própria instituição (dados bibliográficos, pontos de acesso e dados de localização, já definidos nos registros analógicos), a partir do processamento de imagens, o que não requer tratamento adicional dos dados bibliográficos. Para efeito da empiria projetada tomam-se em conta os aspectos teórico-filosóficos constituídos nas pesquisas até então desenvolvidas e a continuidade dos estudos inerentes à proposta de métodos sintáticos e semânticos de registros bibliográficos que sejam aplicáveis aos dados de catalogação na fonte do tipo documental livro.

Este trabalho reveste-se de importância seminal para a área da Ciência da

---

<sup>2</sup> Informação verbal obtida com Luciana Grings, da Biblioteca Nacional, e William Okubo, da Biblioteca Mário de Andrade.



Informação em virtude da existência de poucos trabalhos que abordem esta temática através da modelagem de interpretadores computacionais na forma de reaproveitamento de registros bibliográficos institucionais.

Com base nas considerações apresentadas propõe-se a realização da pesquisa aqui descrita, que tem como objetivo principal discutir o Scan for MARC enquanto interpretador computacional de acesso web para conversão de dados bibliográficos analógicos para o Formato MARC21 Bibliográfico. Entende-se que, assim, seja possível indicar os resultados parciais obtidos na pesquisa e saldar as demandas identificadas nas pesquisas anteriores de modo a disponibilizar o artefato para avaliação, ensaios, testes e uso pelas comunidades de prática.

Relatos de diagnósticos, experiências e procedimentos metodológicos utilizados em diversas instituições para a realização da conversão de dados estão presentes na literatura brasileira. Entretanto, nota-se que cada instituição que necessite estabelecer o processo de conversão de seus dados bibliográficos para o Formato MARC21 Bibliográfico parte do estudo, conhecimento e delineamento da estrutura do seu banco de dados da base a ser convertida.

Oliveira et al. (1998) e Côrte et al. (2002) apresentam a expressão conversão retrospectiva para designar o processo de mudança do suporte no qual um registro bibliográfico está inscrito ou, ainda, o processo de conversão de registros que ainda estão disponíveis para consulta manual, por meio de fichas, para o processo automatizado, em meio magnético, legível por máquina.

Oliveira et al. (1998, p. 42) ressaltam que o método de conversão retrospectiva a ser adotado deve adequar-se “às características e necessidades das bibliotecas”, de modo a garantir “prazos de execução razoáveis, custo baixo e [...] a qualidade das informações.” Côrte et al. (1999) destacam como opções para as bibliotecas que desejam incluir seus registros em fichas em um catálogo automatizado que procedam a conversão retrospectiva, para o que recorrem à digitação de todos os registros no novo sistema, à digitalização destas fichas, ou à busca do registro em outras bases. No Brasil, segundo Oliveira et al. (1998, p. 41), bibliotecários perceberam a necessidade de





conversão retrospectiva de seus acervos na década de 1980 ao se depararem com “dificuldades em registrar, no formato eletrônico, grande parte da produção intelectual da humanidade depositada nas bibliotecas e livrarias.”

Garrido Arilla (2001) afirma que o trabalho de catalogação em ações de conversão retrospectiva é quase nulo, uma vez que o registro já está feito. Entretanto, vários trabalhos técnicos tornam-se necessários. Um deles é o confronto dos documentos com a base de dados fonte e a captura dos registros correspondentes à chave de busca que seja mais apropriada à revocação com maior índice de precisão. Outra atividade é a análise do grau de adequação dos registros com a base de dados destino e, conseqüentemente, a sua correção, momento em que é necessário suprimir alguns campos, modificar o conteúdo de outros, além de acrescentar informação local, etc. Souza e Mostafa (1999) também indicam que a conversão retrospectiva favorece a descrição do recurso, embora problemas quanto à definição dos pontos de acesso possam estar presentes.

Perez e Lima (2002), por sua vez, destacam que a tarefa de conversão retrospectiva é onerosa e intensa e requer a observância de dois aspectos fundamentais na definição do método a ser adotado: um deles recai no custo que se torna alto em se tratando de tratamento externo dos dados; o outro aborda a necessidade de, dependendo da instituição escolhida como base para a cópia dos registros, fazer a adequação de pontos de acesso de assunto, de notas e outras informações locais. Para o êxito de um projeto de conversão retrospectiva, Hubner (2002) indica a necessidade de identificação de informações do acervo a ser convertido (quantidade, tipo e suporte documental), de aplicação de teste piloto em amostra de registros, de definição de rotinas e fluxos de trabalho, de cálculo de tempo e custos envolvidos, dentre outros. Custos e prazos também são apontados por Oliveira et al. (1998).

Garrido Arilla (2001) estabelece que os objetivos a serem alcançados com a conversão retrospectiva são a criação de um único arquivo legível por máquina do catálogo manual existente, a não duplicação de tarefas, e a possibilidade de se evitar custos de catalogação, uma vez que os registros serão copiados e não digitados.



Compreende-se, assim, que não recorrer à conversão retrospectiva<sup>3</sup> é impensável, haja vista que, por meio dela, é possível intercambiar dados, em meio digital, em redes de cooperação.

Com a cooperação entre bibliotecas estabelecida é possível, por meio de padrões, estabelecer com eficiência a conexão entre bibliotecas, o que evita retrabalho, facilita ações de migração de softwares de gerenciamento das unidades e promove reaproveitamento de registros bibliográficos. A catalogação cooperativa, uma das ações resultado da cooperação entre bibliotecas reservada às unidades de informação que contam com gerenciamento automatizado de seu acervo, é resultado da colaboração em catálogos coletivos online, que conta, por sua vez, com fundos de centros participantes e eleva o conceito de dados de dados cooperativa. Um requisito importante para a catalogação cooperativa diz respeito à estrutura e ao formato de descrição do conteúdo na representação dos recursos informacionais, uma vez que isso evita a redigitação dos dados. Para tanto, é requerida a garantia da base tecnológica e metodológica fornecida pela adoção de padrões, que, por sua natureza, promovem compatibilidade. Marcondes (1997), Côrte et al. (1999), Café, Santos e Macedo (2001), Paranhos (2004) e Dziekaniak et al. (2008) indicam que, para fins de intercâmbio de registros bibliográficos, os padrões a serem adotados, indicados por privilegiarem o compartilhamento de dados, envolvem a adoção de um padrão de representação do conteúdo do documento (AACR2r, por exemplo), de um padrão de estrutura de um registro bibliográfico (Formato MARC 21), de um padrão com estrutura interpretável por computadores para a comunicação entre diferentes sistemas (ISO 2709), e de um protocolo que autentica a comunicação em uma estrutura cliente-servidor, com fins de

---

<sup>3</sup> Leituras relacionadas à área em tela e produzidas pelos mais diversos autores podem e devem ser realizadas. Indica-se como leitura complementar à citada nesta pesquisa e como relevante para a compreensão do cenário em que se insere a conversão de registros bibliográficos e, mais especificamente, a metodologia adotada nesta atividade, a leitura de Paranhos (2004), Castro e Santos (2005), Silva e Favaretto (2005), Silva, Alves e Viapiana (2008), Zafalon e Santos (2010), Zafalon e Santos (2012a, 2012b, 2012c), Zafalon (2013), Assumpção e Santos (2014), Zafalon (2014), Zafalon, Santos e Silva (2015a, 2015b).



identificação e recuperação de registros bibliográficos (Z39.50).

A conversão retrospectiva assume-se como requisito essencial para o estabelecimento da cooperação e o intercâmbio de registros bibliográficos para aquelas bibliotecas que ainda não fazem uso de sistemas automatizados de gerenciamento de bibliotecas, ou, minimamente, para a disponibilização de catálogo de acesso público online. Cunha (1979) e Ferreira et al. (1979) afirmam que a catalogação cooperativa no Brasil iniciou-se em 1942, ganhou destaque entre os anos de 1947 e 1952 e, a partir de 1954, tornou-se fundamental para o estabelecimento de projetos e serviços cooperativos nacionais, apesar de serem observados os princípios de catalogação cooperativa desde a década de 40. Entretanto, apesar de passados tantos anos, registros bibliográficos em meios analógicos ainda existem (tal qual indicado no início deste texto), e, mais que isso, o retrabalho em processos de conversão ainda se faz presente.

## MÉTODO DA PESQUISA

Como procedimentos metodológicos destaca-se a afirmação de Marconi e Lakatos (2001, p. 83) de que “não há ciência sem o emprego de métodos científicos” e que o “método é o conjunto das atividades sistemáticas e racionais que, com maior segurança e economia, permite alcançar o objetivo”. Desse modo, o percurso metodológico definido nesta pesquisa baseia-se nas considerações de Gonsalves (2011) e de Silva e Menezes (2005).

A pesquisa apresenta abordagem qualitativa por assumir uma relação dinâmica com o mundo real, por meio da interpretação e atribuição de significados aos fenômenos estudados. Segundo seus objetivos, a pesquisa possui caráter exploratório, uma vez que “se caracteriza pelo desenvolvimento e esclarecimento de ideias, com objetivo de oferecer uma visão panorâmica, uma primeira aproximação a um determinado fenômeno que é pouco explorado” e por oferecer, por meio da pesquisa bibliográfica, “dados elementares que dão suporte para a realização de estudos mais aprofundados sobre o tema,” (GONSALVES, 2011, p. 67).



Por considerar, ainda, os objetivos, a pesquisa encarrega-se do caráter descritivo, por “descrever as características de um objeto de estudo.” (op. cit., p. 68). O caráter experimental está presente por referir-se, também, a “um fenômeno que é reproduzido de forma controlada, submetendo os fatos à experimentação (verificação), buscando, a partir daí, evidenciar as relações entre os fatos e as teorias.” (op. cit., p. 69).

Por configurar-se como pesquisa de cunho teórico e prático propõe-se a abordagem de temas relacionados à Ciência da Informação, em especial estudos inerentes à representação documental, aos catálogos, à conversão e ao compartilhamento de registros bibliográficos; à Ciência da Computação, em decorrência das pesquisas voltadas ao processamento de imagens, às técnicas de reconhecimento óptico de caracteres, às tecnologias que envolvem o desenvolvimento de interpretadores sintáticos bem como o desenvolvimento de inferências semânticas e o desenvolvimento de ambientes web para processamento de dados; e à Linguística, diante dos estudos sintáticos e semânticos que embasam aspectos teóricos para análise sintática e semântica presentes nos estudos linguísticos da comunicação humana e da concepção estruturalista da língua.

Para tanto, propõe-se procedimentos metodológicos de análise exploratória e crítica de textos relevantes ao tema da pesquisa. A análise de textos apresenta caráter exploratório, de modo a contextualizar as características sobre o objeto em estudo apresentado nos objetivos propostos. A análise crítica da literatura cria oportunidades para elaboração e contribuição para o conhecimento teórico da área de Ciência da Informação. Aspectos tecnológicos estão embutidos nos procedimentos que envolvem a análise da performance do interpretador computacional.

Desse modo, apresentam-se os resultados das interfaces de processamento do Scan for MARC.





## DISCUSSÃO E ANÁLISE DOS RESULTADOS

Com base na proposta teórica-conceitual de análise sintática e semântica do registro bibliográfico<sup>4</sup>, pilar para a conversão retrospectiva proposta como solução para o problema em questão – adequações no registro bibliográfico do sistema-base da catalogação cooperativa, de modo que o novo registro deva refletir informações locais – , apresenta-se o Scan for MARC.

Os pressupostos teóricos para a análise sintática e semântica de registros bibliográficos, subsídio para o desenvolvimento do Scan for MARC, envolvem:

- a) em registros bibliográficos, a sintaxe é identificada pela pontuação, indicada em padrões de estrutura de metadados descritivos, nesse caso a ISBD e as regras de pontuação das AACR2r. Estudam-se a forma, o arranjo e a disposição de cada elemento descrito na representação do recurso informacional, e, com base em Saussure (2010), compreende-se que a sintaxe estabelece as relações formais e síncronas entre *significado*, *significante*, *obra* e *manifestação*;
- b) em registros bibliográficos a semântica, estabelecida a partir da teoria dos níveis do estudo estruturalista dos registros bibliográficos, com base em Hjelmslev (1991), ocorre em três momentos:
  - *subsemântico*: estudo da teoria das unidades semânticas ou dos elementos semanticamente relacionados presentes em um registro bibliográfico; significado estabelecido entre cada um dos elementos de um registro e os valores de representação que assume;
  - *semântico*: descreve as relações estruturais estabelecidas entre o registro bibliográfico e o recurso informacional;
  - *supersemântico*: relação semântica entre os vários registros bibliográficos representados em um catálogo, por meio de identidade de significados (semelhanças) e de oposição de significados (diferenças). (ZAFALON, 2014).

O Scan for MARC foi validado em ambiente de acesso local, com as ações em linha de comando e com processamento com interação humana na captura das imagens (ZAFALON, 2012a; 2013; 2014). Nesta primeira fase foi possível: definir funcionalidade sintática e semântica; proceder captura de imagens em meio eletrônico;

---

<sup>4</sup> A proposta teórico-conceitual que subsidia a conversão retrospectiva a partir de análise sintática e semântica de registros bibliográficos pode ser avaliada em Zafalon (2013, 2014).



avaliar funcionalidades de softwares OCR; testar aplicabilidade de modelos; realizar processamento de imagens; e interpretar dados em linha de comando. A figura 1 indica os resultados obtidos nesta fase.

Figura 1 – Resultados da fase 1 do Scan for MARC (linha de comando)

<pre>830.1 Dubell, Richard D111b A bíblia do diabo : romance histórico / Richard Dubell ; tradução Claudia Abeling. - São Paulo : Planeta do Brasil, 2011. 512 p.  Título do original: Die Terfelsbibel.  1. ROMANCE HISTÓRICO 2. LITERATURA ALEMÃ I. Abeling, Claudia. II. Título.</pre>	<pre>021.3 Leitura e escrita de adolescentes na internet e na L21 escola / Organização Maria Teresa de Assunção 2.ed. Freitas, Sérgio Roberto Costa. - 2. ed. - Belo Horizonte : Autêntica, 2006. 138 p. - (Coleção leitura, escrita e oralidade)  ISBN 85-7526-156-8.  1. TECNOLOGIA DA INFORMAÇÃO 2. LEITURA I. Freitas, Maria Teresa de Assunção, org. II. Costa, Sérgio Ro- berto, org. III. Série.</pre>
<pre>090 \$a830.1 \$bD111b 1001 \$aDubell, Richard 24512\$aA bíblia do diabo \$bromance histórico \$cRichard Dubell ; tradução Claudia Abeling 260 \$aSão Paulo \$bPlaneta do Brasil \$c2011 300 \$a512 p. 500 \$aTítulo do original: Die Terfelsbibel 65014\$aROMANCE HISTÓRICO 65024\$aLITERATURA ALEMÃ 7001 \$aAbeling, Claudia</pre>	<pre>020 \$a8575261568 090 \$a021.3 \$bL21 \$c2.ed. 24500\$aLeitura e escrita de adolescentes na internet e na escola \$cOrganização Maria Teresa de Assunção Freitas, Sérgio Roberto Costa 250 \$a2. ed. 260 \$aBelo Horizonte \$bAutêntica \$c2006 300 \$a138 p. 4900 \$aColeção leitura, escrita e oralidade 65014\$aTECNOLOGIA DA INFORMAÇÃO 65024\$aLEITURA 7001 \$aFreitas, Maria Teresa de Assunção \$eorg. 7001 \$aCosta, Sérgio Roberto \$eorg</pre>

Fonte: Elaboração própria.

A partir dos resultados obtidos com os testes realizados com o *script* em linha de comando foi possível utilizá-lo como base para o desenvolvimento do interpretador em ambiente de acesso web (fase dois), com interface disponível para acesso no endereço [www.scanformarc.ufscar.br](http://www.scanformarc.ufscar.br)<sup>5</sup>.

Esta fase consagrou-se por ações que envolviam o texto como input e a codificação ISO, que pode ser observada na figura 2.

<sup>5</sup> Neste endereço é possível conferir duas versões do interpretador: uma delas para a digitação ou colagem do texto, utilizada para verificação da adequabilidade de script (fase dois); outra, com a integração do script com o OCR, para o qual adotou-se o Tesseract (fase três).



Figura 2 – Resultados da fase 2 do Scan for MARC (*web*)

### Scan for Marc

<p>631.521 Toledo, Francisco Ferraz de T649m Manual das sementes : tecnologia da produção / Francisco Ferraz de Toledo, Júlio Marcos Filho. - São Paulo : Agronômica Ceres, 1977. 224 p.  Inclui bibliografia.  1. SEMENTES 2. PRODUÇÃO I. Marcos Filho, Júlio. II. Título.</p>	<p>NOTE: using character set ISO-8859-1  ===== MARC ===== 090 \$a631.521 \$bT649m 1001 \$aToledo, Francisco Ferraz de 24510\$aManual das sementes \$btecnologia da produção \$cFrancisco Ferraz de Toledo, Júlio Marcos Filho 260 \$aSão Paulo \$bAgronômica Ceres \$c1977 300 \$a224 p. 504 \$aInclui bibliografia 65014\$aSEMENTES 65024\$aPRODUÇÃO 7001 \$aMarcos Filho, Júlio =====</p>
<p><a href="#">Clear</a></p>	<p><a href="#">Clear</a></p>

ScanForMARC >>

Fonte: Elaboração própria.

O objetivo definido para esta fase de testes era o de validar a aplicação do *script* na versão *web*, sem que houvesse preocupação com a decodificação da imagem, possível com o OCR embutido no *script* de interpretação. A partir dos resultados obtidos nesta fase optou-se pela submissão de projeto ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) com a finalidade de conseguir aporte financeiro para a continuidade da pesquisa.

A terceira fase de desenvolvimento do Scan for MARC iniciou-se em janeiro de 2015 e foi possível obter respostas positivas quanto [1] à identificação de elementos de representação documental presentes em registros bibliográficos de livros, com as mais diversas complexidades e níveis de granularidade; [2] ao desenvolvimento de modelos de descrição, baseados em roteiros de representação desenvolvidos a partir de padrões de estrutura de metadados descritivos, de padrões de conteúdo e de padrões de visualização de registros bibliográficos; [3] à aplicação de modelos de descrição em interpretador computacional de conversão de registros analógicos para o Formato MARC21 Bibliográfico. Resultados preliminares e, portanto, parciais, quanto à viabilização da conjugação integrada das fases de reprodução e de representação,







A continuidade do projeto ao longo deste ano (2016) permitirá a avaliação dos objetivos 5 a 11 e sua realização. Apesar de não terem sido definidos objetivos específicos para a interpretação de conteúdo em idiomas específicos, foi possível conjugar esforços para a interpretação de registros em português, inglês, espanhol, francês e alemão. Além disso, conquanto não previstos, procedeu-se a aplicação de filtro nas imagens bem como a interpretação de dados codificados em ASCII, UTF8 e WIN1252.

## CONSIDERAÇÕES FINAIS

Tendo em vista os resultados anteriores, obtidos em testes desenvolvidos em ambiente local (com processamento dos dados via linha de comando), embora considerados iniciais, destaca-se que serviram como mote para a continuidade do desenvolvimento da pesquisa, em ambiente *web*, a qual apresentou os resultados parciais aqui indicados.

Tendo em vista a consolidação dos objetivos definidos para a continuidade da pesquisa, espera-se, ao longo de 2016, o cumprimento das metas na quarta fase: [1] desenvolver interface gráfica que permita upload e download de registros bem como registro de avaliações pelas comunidades de prática; [2] identificar padrões para a interpretação de dados de classificação de assunto (CDU); [3] prover interpretação do conteúdo de termos qualificadores (subdivisão de forma, subdivisão cronológica, subdivisão geográfica e subdivisão geral) nos pontos de acesso de assunto tópico; [4] prover interpretação de pontos de acessos secundários de assuntos para nome pessoal, institucional, evento e título uniforme; [5] prover interpretação de pontos de acessos principais para instituições, eventos e título uniforme; [6] desenvolver dicionários semânticos para auxílio nos processos de inferência de pontos de acesso e seus termos qualificadores; [7] implementar processo de teste de consistência do arquivo convertido, uma vez que isso afere caráter qualitativo ao registro bibliográfico gerado. Como metas secundárias serão necessários testes de resolução de imagens, testes de processamento



do software (a ser desenvolvido em ações coordenadas junto à UFSCar, à UNESP e à UDESC) e, por fim, testes pela comunidade.

---

### **Scan for MARC: conversion records in cards to the MARC21 Format for Bibliographic Data**

**Abstract:** Discusses conversion of bibliographic records to be used by institutions that do not provide all data in their collections online public access catalogs. Despite the possibility of using the retrospective conversion processes of bibliographic records to pay off such a gap, technical work become necessary, such as adequacy of subject access points, notes and other local information on which incur correction, suppression fields, modifying content and addition of local information. It presents the retrospective conversion of bibliographic data from the image processing, which considers the full exploitation of the bibliographic record of the institution (bibliographic data, access points and location data, as defined in analog records). It is proposed as a main objective to present the Scan for MARC, a computational interpreter of web access for converting analog bibliographic data for the MARC21 Format Library. It resorts to a qualitative approach with exploratory and descriptive, and bibliographic research. The results show contributions in scientific, before the development of syntactic and semantic aspects of bibliographic records; in technological terms, given the development of the Scan for MARC as a product to improve conversion processes analog bibliographic data; social, inherent in sharing records and their methodologies, the expansion of access to metadata and to interoperability between managers of library systems, factors that contribute to the discussion and socialization of scientific / technological knowledge.

**Keywords:** Scan for MARC. Retrospective conversion. Cataloging. Library automation. Online Public Access Catalog.

### **REFERÊNCIAS**

ASSUMPCÃO, F. S.; SANTOS, P. L. V. A. C. Conversão de registros em XML para MARC 21: um modelo baseado em XSLT. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA D INFORMAÇÃO, 15., 2014, Belo Horizonte. *Anais...* Belo Horizonte: ECI/UFMG, 2014. p. 3802-3821.



- CAFÉ, L.; SANTOS, C.; MACEDO, F. Proposta de um método para escolha de software de automação de bibliotecas. *Ciência da Informação*, Brasília, v. 30, n. 2, p. 70-79, maio/ago. 2001.
- CASTRO, F. F.; SANTOS, P. L. V. A. C. Conversão retrospectiva de registros bibliográficos. In: FUJITA, M. S. L. (Org.) et. al. *A dimensão social da biblioteca digital na organização e acesso ao conhecimento: aspectos teóricos e aplicados*. São Paulo: DT/SIBi : IBICT, 2005. v. 2, p. 63-87.
- CÔRTE, A. R. et al. Automação de bibliotecas e centros de documentação: o processo de avaliação e seleção de softwares. *Ciência da Informação*, Brasília, v. 28, n. 3, p. 241-256, set./dez. 1999.
- CÔRTE, A. R. et al. *Avaliação de softwares para bibliotecas e arquivos*. 2. ed. rev. ampl. São Paulo: Polis, 2002.
- CUNHA, M. L. M. ISBD: origem, evolução e aceitação. *Revista Brasileira de Biblioteconomia e Documentação*, v. 12, n. 1/2, p. 7-14, jan./jun. 1979.
- DZIEKANIAK, G. V. et al. Uso do padrão MARC em bibliotecas universitárias da Região Sul do Brasil. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, n. 26, jul./dez. 2008.
- FERREIRA, J. R. et al. Redes nacionais de informação, catalogação na fonte e outras experiências. *Revista Brasileira de Biblioteconomia e Documentação*, v. 12, n. 1/2, p. 67-88, jan./jun. 1979.
- GARRIDO ARILLA, M. R. Procedimientos automáticos de creación y transformación de los registros bibliográficos. *Anales de Documentación*, n. 4, p. 127-137, 2001.
- GONSALVES, E. P. *Conversas sobre iniciação à pesquisa científica*. 5. ed. rev. e ampl. Campinas: Alínea, 2011.
- HJELMSLEV, L. *Ensaio lingüísticos*. São Paulo: Perspectivas, 1991.
- HÜBNER, E. Conversão retrospectiva de registros bibliográficos. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 13., 2002, Recife. *Anais...* Recife: UFPE, 2002.
- LANCASTER, F. W. Ameaça ou oportunidade? O futuro dos serviços bibliotecários à luz das inovações tecnológicas. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 23, n. 1, p. 7-27, jan./jul. 1994.
- MARCONDES, C. H. Automação de bibliotecas. In: CURSO Aplicação de tecnologia ao desenvolvimento de bibliotecas. Rio de Janeiro: FGV, 1997.
- MARCONI, M. A.; LAKATOS, E. M. *Fundamentos de metodologia científica*. 4. ed.



rev. e ampl. São Paulo: Atlas, 2001.

OLIVEIRA, N. M. et al. Compact disc cataloging – CatCD: análise de um instrumento para conversão retrospectiva no Sistema de Bibliotecas da UNICAMP. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 3, n. 1, p. 41-46, jan./jun. 1998.

PARANHOS, W. M. M. R. Fragmentos metodológicos para projetos e execução de gestão informatizada de coleções de documentos e serviços em bibliotecas. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, n. esp., jul./dez. 2004.

PEREZ, D. R.; LIMA, P. O projeto de conversão retrospectiva de registros bibliográficos: uma experiência do sistema de bibliotecas da Pontifícia Universidade Católica do Rio de Janeiro. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 12., 2002, Recife. *Anais...* [s. l.: s.n.], 2002.

SAUSSURE, F. *Curso de lingüística geral*. São Paulo: Cultrix, 2010.

SILVA, F. C. C.; FAVARETTO, B. Uso de softwares para o gerenciamento de bibliotecas: um estudo de caso da migração do sistema Aleph para o sistema Pergamum na Universidade de Santa Cruz do Sul. *Ciência da Informação*, Brasília, v. 34, n. 2, p. 105-111, maio/ago. 2005.

SILVA, F. C. L.; ALVEZ, G.; VIAPIANA, N. Informatização da rede de bibliotecas da Secretaria Municipal de Educação de Florianópolis. *Revista ACB: Biblioteconomia em Santa Catarina*, Florianópolis, v. 13, n. 1, p. 211-222, jan./jun. 2008.

SOUZA, T. B.; MOSTAFA, S. P. Catalogação cooperativa na rede BIBLIODATA/CALCO: a questão da repetitividade dos títulos no catálogo coletivo. *Transinformação*, Campinas, v. 11, n. 2, p. 127-133, maio/ago. 1999.

ZAFALON, Z. R. *Scan for MARC: conversão de registros bibliográficos analógicos para o Formato MARC21 Bibliográfico*. São Paulo: Ed. UNESP, 2014.

ZAFALON, Z. R. Scan for MARC: sintaxe e semântica de registros bibliográficos na conversão de dados analógicos para o Formato MARC21 Bibliográfico. *Revista CRB-8 Digital*, São Paulo, v. 6, p. 20-32, 2013.

ZAFALON, Z. R., SANTOS, P. L. V. A. C. Aplicações sintáticas e semânticas na conversão de registros bibliográficos ao formato MARC21 Bibliográfico In: ENCUESTRO INTERNACIONAL DE CATALOGADORES, 7., 2012, Buenos Aires; ENCUESTRO NACIONAL DE CATALOGADORES, 3., 2012, Buenos Aires. *Estándares y procedimientos para la organización de la información*. [s.l.: s.n.], 2012b. Disponível em: <http://www.bn.gov.ar/descargas/catalogadores/encuentro2011/ponencia-24-C-Zafalon-Santos.pdf>.

ZAFALON, Z. R., SANTOS, P. L. V. A. C. Conversión de registros bibliográficos al





Formato MARC21 Bibliográfico a partir del análisis sintáctico y semántico de registros descritos según las AACR2r y el RDA In: ENCUENTRO DE CATALOGACIÓN Y METADATOS, 4., 2010, Ciudad de México. *Memoria*. Ciudad de México: UNAM, 2010. p.127-148.

ZAFALON, Z. R., SANTOS, P. L. V. A. C. Sintaxe e semântica de registros bibliográficos: princípios para a conversão de registros analógicos para o Formato MARC21 Bibliográfico: o Scan for MARC In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 13., 2012, Rio de Janeiro. *Anais digitais*. Rio de Janeiro: ICICT/Fiocruz, 2012a.

ZAFALON, Z. R., SANTOS, P. L. V. A. C. Sintaxe e semântica de registros bibliográficos: princípios para a conversão de registros analógicos para o Formato MARC21 Bibliográfico: Scan for MARC. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 5, p. 1-23, 2012c.

ZAFALON, Z. R., SANTOS, P. L. V. A. C., SILVA, J. Scan for MARC: conversor de dados analógicos para o formato MARC21 bibliográfico. In: ENCUENTRO IBÉRICO EDICIC, 7., 2015, Madrid. *Desafíos y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas...* Madrid: Universidad Complutense de Madrid, 2015a.

ZAFALON, Z. R., SANTOS, P. L. V. A. C., SILVA, J. Scan for MARC: syntax and semantics of bibliographical records in the conversion of analogic data to the MARC21 Bibliographic Format In: ENCUENTRO DE CATALOGACIÓN Y METADATOS, 7., 2015, México. *Memoria*. 2015b. Disponível em: <http://hdl.handle.net/10760/28162>. Acesso em: 20 dez. 2015. Organizado por Filiberto Felipe Martinez-Arellano.

### **Agência financiadora**

Esta pesquisa é financiada pelo CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico.

---

### Informações dos autores

**Zaira Regina Zafalon**  
UFSCar  
Email: [zzafalon@gmail.com](mailto:zzafalon@gmail.com)



**Plácida Leopoldina Ventura Amorim da Costa Santos**

UNESP

Email: placida@marilia.unesp.br

**Ana Maria Pereira, Jairo da Silva**

UDESC

Email: anamariapere@gmail.com

**Jairo da Silva**

UFSCar

Email: jairo.silva@gmail.com

