



SCAN FOR MARC: CONVERSÃO DE REGISTROS EM FICHAS PARA O FORMATO MARC21 BIBLIOGRÁFICO

Zaira Regina Zafalon¹
Plácida L. V. A. da Costa Santos²
Ana Maria Pereira³
Jairo da Silva⁴

Introdução

A pesquisa insere-se no contexto social contemporâneo, no qual se faz uso e aplicação de tecnologias no seu cotidiano o que exige (re)avaliação de processos, produtos e serviços voltados para 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. O principal problema a ser abordado está presente no dilema de que ainda se encontram instituições que não disponibilizam a totalidade de dados bibliográficos de seus acervos em catálogos online de acesso público, ou seja, ainda contam com registros que compõem catálogos bibliográficos impressos em fichas (registros analógicos). Pode-se recorrer a processos de conversão retrospectiva de registros bibliográficos para saldar tal lacuna. Estudos sobre catalogação cooperativa e compartilhamento e conversão de dados bibliográficos são discutidos por Vasconcellos (1996), Oliveira et al. (1998), Garrido Arilla (2001). Entretanto, nos processos tradicionais de conversão retrospectiva, uma vez que registros de catálogos coletivos são utilizados como base para a cópia de registros, vários trabalhos técnicos tornam-se necessários (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). É

¹ Pesquisador do Grupo de Pesquisa Tecnologias em Ambientes Informacionais (GPTAI) e professor na UFSCar, Campus São Carlos. Email: zaira@ufscar.br.

² Pesquisador do GPTAI e professor na UNESP, Campus Marília. Email: placida@marilia.unesp.br.

³ Pesquisador do GPTAI e professor na UDESC. Email: anamariapere@gmail.com.

⁴ Técnico do GPTAI. Email: jairo.silva@gmail.com.

sobre este ponto que se apresenta o propósito 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. Estudos sobre conversão integral de dados bibliográficos por meio do processamento de imagens são discutidos por Zafalon e Santos (2010, 2012a, 2012b, 2012c) e Zafalon (2013, 2014). A pesquisa reveste-se de importância seminal para a área da Ciência da 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. Desse modo, propõe-se como objetivo principal desenvolver o Scan for MARC (SfM), software de conversão de dados bibliográficos em fichas para o Formato MARC21 Bibliográfico, em ambiente web. Entende-se que, assim, seja possível saldar as demandas identificadas nas pesquisas anteriores de modo a disponibilizar o artefato para avaliação, ensaios, testes e uso pelas comunidades de prática.

Método da pesquisa

Quanto aos procedimentos metodológicos, 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 e descritivo. Recorre-se à pesquisa bibliográfica para a compreensão e elaboração do arcabouço teórico, estabelecido a partir de temas relacionados à pesquisa e que integram a Ciência da Informação, a Ciência da Computação e a Linguística. O caráter experimental está presente por referir-se a um fenômeno reproduzido de forma controlada, com experimentação, verificação e estudo das evidências e relações entre fatos e as teorias.

Resultados e Discussão

A partir do desenvolvimento do Scan for MARC em linguagem Perl e aplicação em servidor web foi possível observar a necessidade de ajustes quanto ao software de reconhecimento ótico de caracteres (OCR), quando aplicado em ambiente web. Isso se fez necessário dadas as diferentes aplicações quando se utilizam linguagens para interpretação em linha de comando e em ambiente web. O ambiente de teste está disponível em: www.scanformarc.ufscar.br. Observa-se o alcance parcial dos resultados uma vez que foi possível desenvolver, na versão web, a conjugação integrada das fases de reprodução e de representação, recorrendo-se à adoção de um OCR em linha de comando; o desenvolvimento de interface gráfica; e a integração do script com o OCR, o que exige a adoção de um que seja em linha de comando.

Considerações Finais

A partir da pesquisa desenvolvida observa-se o alcance parcial do objetivo estabelecido, apesar de vislumbrarem-se contribuições no plano científico, tecnológico e social. Para estudos futuros busca-se aprimorar o SfM de modo a contemplar a interpretação de dados de classificação de assunto (CDU); de termos qualificadores (subdivisão de forma, cronológica, geográfica e geral); de pontos de acesso para nome pessoal, institucional, evento e título uniforme, bem como de criação de dicionários semânticos para pontos de acesso e termos qualificadores; e implementação de processo de teste de consistência do arquivo convertido.

Palavras-chave: Conversão bibliográfica. Formato MARC21. Scan for MARC.

Referências

- GARRIDO ARILLA, M. R. Procedimientos automáticos de creación y transformación de los registros bibliográficos. *Anales de Documentación*, n. 4, p. 127-137, 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.
- VASCONCELLOS, P. A. G. Bibliodata/CALCO: informação bibliográfica para o desenvolvimento. *Ciência da Informação*, Brasília, v. 25, n. 3, p. 450-453, set./dez. 1996.
- 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*, 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: VII Encuentro Internacional y III Nacional de Catalogadores, 2012, Buenos Aires. Estándares y procedimientos para la organización de la información, 2012a.
- ZAFALON, Z. R.; SANTOS, P. L. V. A. C. Conversión de registros bibliográficos al Formato MARC21 bibliográfico a partir del análisis sintáctico e semântico de registros descritos según las AACR2r y el RDA. IV Encuentro de Catalogación y Metadatos. *Memória*. México/DF: UNAM/CUIB, 2010.
- 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, 2012b.
- 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, 2012, Rio de Janeiro. *Anais digitais*. Rio de Janeiro: ICICT/Fiocruz, 2012c.

Agência financiadora: CNPq