

APLICAÇÕES SINTÁTICAS E SEMÂNTICAS NA CONVERSÃO DE REGISTROS BIBLIOGRÁFICOS AO FORMATO MARC21 BIBLIOGRÁFICO

Zaira Regina Zafalon

Plácida L. V. A. da C. Santos

UNESP - BRASIL

Introdução

- Uso de novas tecnologias no cotidiano
- Recursos tecnológicos e mediáticos:
 - Produção
 - Organização
 - Distribuição
 - Acesso
- Consequências: contexto social, cultural, educacional e econômico e o crescimento de acervos documentais.

Registros do
Conhecimento

Introdução

- Adequação de P&S às demandas informacionais por diferentes públicos.
- Facilidades de compartilhamento de registros bibliográficos entre as unidades de informação.
- Unidades de informação beneficiam-se de compartilhamento de registros e de interoperabilidade de bases de dados.

Introdução

- Diferentes unidades de informação procuram adequar-se aos padrões internacionais de descrição e compartilhamento de dados bibliográficos (AACR2r e RDA).
- Na literatura: o processo de conversão de dados bibliográficos para o formato MARC21 Bibliográfico parte do estudo, conhecimento e delineamento da estrutura dos dados na base a ser convertida.

Representação bibliográfica e comunicação para e com usuários

Objetivo de representação bibliográfica é estabelecer a comunicação entre os recursos informacionais disponíveis e as necessidades de seus usuários.

Para a operacionalização da comunicação é necessária a representação, baseada em normas, padrões e recursos.

Cooperação e compartilhamento de registros bibliográficos

- 200 a.C. – Alexandria Library
- Séc. XIII – monastérios
- Século XVII – reunião de catálogos
- 1885 – compilação de periódicos – Henry Bolton
- Início do Séc. XX – projetos de catalogação pela Library of Congress
- 1886 – publicação de Dewey sobre cooperação entre bibliotecas

Catálogo e automação

- Pré-requisito para outras formas de cooperação e supõe esforço para a normalização dos dados.
- Catalogação cooperativa – não catalogar o que já está catalogado em algum ponto da rede.

Catálogo e automação

- Processo comunicativo na catalogação cooperativa:
 - Emissor – agência de catalogação cooperativa
 - Mensagem – descrição bibliográfica
 - Canal – catálogo coletivo online
 - Receptor – usuários da agência ou rede bibliotecária

Garrido Arilla, 2001

Compartilhamento de registros

- Padrão de representação dos documentos
 - AACR2
- Formato de catalogação legível por máquina
 - MARC21
- Formato de intercâmbio
 - ISO 2709
- Protocolo de comunicação
 - Z39.50

Compartilhamento de registros

BRASIL

- Grande parcela das unidades de informação não se utilizam do MARC21 para compartilhamento de registros bibliográficos e para estabelecimento de redes de cooperação
- Região Sul, cerca de 63% se utilizam do MARC21

Compartilhamento de registros

BRASIL

Por quê?

- o desconhecimento dos padrões ou das vantagens em adotá-los;
- as limitações oriundas de infra-estrutura tecnológica;
- as restrições orçamentárias para o tratamento dos dados;
- a falta de habilidades de gestão para o gerenciamento de diferenciados recursos necessários ao processo de mudança ou implantação de sistemas automatizados;

Compartilhamento de registros

BRASIL

Por quê?

- a inexperiência no processo de seleção de software de gerenciamento que atenda aos requisitos de compartilhamento de dados e uso do formato MARC21 no delineamento das bases de dados;
- o desconhecimento dos processos necessários tanto para a conversão quanto para a migração de dados bibliográficos, apesar de se ter conhecimento da existência de padrões e das vantagens na adoção.se utilizam do MARC21 para compartilhamento de registros bibliográficos e para estabelecimento de redes de cooperação.

Proposta ...

- Metodologia de conversão de registros bibliográficos para o formato MARC21 Bibliográfico por meio de uma mesma metodologia aplicável a diferentes bases;
- Possibilidade de conversão a partir do conteúdo sintático, descrito pelas AACR2r ou RDA, e semântico do registro bibliográfico, pelas AACR2r e pelos FRBR e os FRAD;
- identificação de técnica computacional necessária para o entendimento do conteúdo sintático e semântico do registro bibliográfico.

Pressuposto ...

Há outra forma de converter dados bibliográficos desconsiderando-se a estrutura da base de dados

Processo adotado atualmente: estudo da estrutura da base de dados para:

- [1] o reconhecimento das etiquetas de campo (TAGs) utilizadas,
- a análise de equivalências entre elas e o formato MARC21 Bibliográfico, e
- o rearranjo dos dados no padrão MARC21 Bibliográfico.

Estrutura da base de dados 'X'

CLASS	AUTOR	TITUL	ANO AQ	AQUIS	LOCAL	EDITO	ANOPB	PAGIN	CDD	IDIOM	TIPO	BIBLI	ASSUN
270.02	Pierini, Franco	A idade antiga: curso de	2004	Compra	São Paulo	Paulus	1998	249 p.	21	Port.	Livro	Bibliografia	Curso de história e
221.93	Lenhardt, Pierre; Collin, M	A Torah oral dos fariseu	2004	Compra	São Paulo	Paulus	1997	155 p.	21	Port.	Livro		Documentos do m
262.01	Kehl, Medard	A igreja: uma eclesiolog	2004	Compra	São Paulo	Loyola	1997	414 p.	21	Port.	Livro	Bibliografia	Eclesiologia
225.83	Arens, Eduardo	Ásia menor nos tempos	2004	Compra	São Paulo	Paulus	1998	209 p.	21	Port.	Livro		Sociologia bíblica
264.025		Sagrada congregação	2001	Compra	São Paulo	Paulus	2001	308 p.	22	Port.	Livro	Bibliografia	Rituais \$ Iniciação
261.805	Hoornaert, Eduardo	A memória do povo cris	1986	Doação	Petrópolis	Vozes	1986	364 p.	22	Port.	Livro		Cristianismo primit
221.9	Miranda, Evaristo Eduard	Sábios fariseus: repara	2004	Compra	São Paulo	Loyola	2001	239 p.	23	Port.	Livro		Judaísmo
220.63	Trahalla Berrara, Julia C	A Bíblia judaica e a Bibl	2004	Compra	Petrópolis	Vozes	1999	741 p.	22	Port.	Livro		Bíblia judaica \$ Cr

Campos da estrutura da base de dados 'X'

Etiquetas de campo constantes da estrutura da base de dados 'X' são:

- CLASS (para o número de Classificação),
- AUTOR (para o nome do(s) autor(es)),
- TITUL (para o título),
- ANOAQ (para o ano de aquisição),
- AQUIS (para a forma de aquisição),
- LOCAL (para o local de publicação),
- EDITO (para a editora),
- ANOPB (para o ano de publicação),
- PAGIN (para a descrição das páginas)

Campos da estrutura da base de dados 'X'

Etiquetas de campo constantes da estrutura da base de dados 'X' são:

- CDDDED (para a identificação da edição da Tabela de Classificação Decimal de Dewey),
- IDIOM (para o idioma do item),
- TIPO (para o tipo de material),
- BIBLI (para o tipo de bibliografia do curso) e
- ASSUN (para identificação da temática da obra).

Estrutura da base de dados 'Y'

Assunto	Autor(a)	Classificação	Not. Autor	Série	Editora	Curso	Título
Antropologia - Antropologia relig	Mancini, Roberto	233	M237e	Coleção 1	Paulinas		Existência e Gratuidade
Antropologia - Antropologia teol	Fabbri, Enrique	233	F113f	Conviver	Paulinas		Família
Antropologia teológica - Cultura	Gritti, Jules	233	G892e	Crer e compreender	Paulinas		Expressão da fé nas cu
Bíblia - arqueologia bíblica - Auen	Rolla, Armando	22.08	R655a	Coleção 1	Paulinas		A Bíblia e as últimas des
Bíblia - Povo de Deus: história	CRB	22.096	C1b	Coleção 1	Loyola		A Bíblia na formação
Catolicismo Negro - Teologia Neg	Da Silva, Antônio Aparecido	239.8	S578e	Atabaque- cultura negra e teologia	Paulinas		Existe un pensar teológi
Credo - Santo Tomás de Aquino	Tomás de Aquino	276.12	T611e	Coleção 1	Presença		Exposição sobre o Cred
Deus - Cristianismo	Catão, Francisco	231	C354f	Iniciação à Teologia	Paulinas		Falar de Deus
Deus - Deus - idéia de Deus - pi	VV.AA. Frei Betto D. Baggio, Hi	231.74	VV.AA.	Teologia	Vozes		Experimentar Deus hoje
Ética familiar - Destacel familiar - E	Solomão, Janice	241.04-173	S17r	Coleção 1	Paulinas		Família

Campos da estrutura da base de dados 'Y'

Etiquetas de campo constantes da estrutura da base de dados 'Y':

- Assunto,
- Autor(a),
- Classificação,
- Not. Autor,
- Série,
- Editora,
- Curso e
- Título.

Variação na estrutura das bases de dados

Sem padronização para as etiquetas de campos

Sem possibilidade de compartilhamento de registros a partir da estrutura da base de dados

Proposta de análise sintática em
semântica em registros bibliográficos
para conversão

Sintaxe – arranjo, disposição, ordenação
e organização.

Sintaxe do registro bibliográfico:

ordem dos elementos dispostos para a
representação dos registros do conhecimento. É
parte da Ciência da Informação voltada ao estudo
da forma, arranjo e disposição em que cada
elemento deva ser descrito quando da
representação do registro do conhecimento.

Determina relações formais entre as cada uma das
partes do registro representado.

Sintaxe do registro bibliográfico

- *Santos, Plácida Leopoldina Ventura Amorim da Costa*
 - são formas de apresentação dos elementos de responsabilidade intelectual por uma obra como pontos de acesso ao documento, dadas pela Regras 21.1, 22.1, 22.4 e 22.5 das AACR2r;
- *Catálogo : trajetória para um código internacional*
 - é a forma de apresentação do elemento título e de outras informações sobre o título de uma obra, dada pelas Regras 1.1B, 1.1E, 2.1A1, 2.1B, 2.1E das AACR2r;
- *Niterói : Intertexto, 2009*
 - são formas de apresentação dos elementos local de publicação, nome do editor e data de publicação de uma obra, dadas pelas Regras 1.4B, 1.4C, 1.4D, 1.4F, 2.4A1, 2.4B, 2.4C, 2.4D e 2.4F das AACR2r.

Proposta de análise sintática em semântica em registros bibliográficos para conversão

Semântica – significação, sentido,
significado.

Semântica do registro bibliográfico:

significado estabelecido entre cada um dos
elementos representados e o documento
representado, propondo-se um conteúdo
semântico do registro bibliográfico, quanto ao
significado entre o elemento definidor e o valor
dado para expressar este elemento.

Proposta de análise sintática em semântica em registros bibliográficos para conversão

Em outras palavras:

Semântica do registro bibliográfico:

o registro bibliográfico de um documento só pode assumir o significado de representação daquele documento dado o sentido que pode ser estabelecido entre representado e representante;

e, também, o significado entre o elemento a ser representado e o seu valor (dado, por sua vez, pela sintaxe).

Semântica em registros bibliográficos



Catalogação : trajetória para um código internacional

Estes elementos assumem, no contexto de um registro bibliográfico, a representação que dá o significado de título e de outras informações para o título da obra.

Semântica em registros bibliográficos



- *Niterói : Intertexto, 2009*

Estes elementos, por sua vez, assumem, no contexto de um registro bibliográfico, a representação que dá o significado de elementos como local de publicação, nome do editor e data de publicação de uma obra.

Semântica em registros bibliográficos

Explicita-se:

[1] a semântica entre os elementos de uma obra (título, outras informações sobre o título, local de publicação, nome do editor e data de publicação) e o seu valor (o seu conteúdo designativo e representativo da obra)

Semântica em registros bibliográficos

Explicita-se:

[2] a semântica em um registro bibliográfico, na qual, a representação bibliográfica, para que assuma a sua vertente representacional e significativa de um representado, apresente relações entre os valores dos elementos, ou seja, no exemplo acima, a obra intitulada *Catálogo : trajetória para um código internacional* apresenta, relação intrínseca com a sua manifestação, dada pelos elementos da publicação, tais como, *Niterói : Intertexto, 2009.*

Sintaxe e semântica em registros bibliográficos e o Formato MARC21

- 100 1# \$aSantos, Plácida Leopoldina Ventura Amorim da Costa
- 245 10 \$aCatalogação \$btrajetória para um código internacional
- 260 ## \$aNiterói \$bIntertexto \$c2009

A sintaxe em cada conteúdo de TAG do Formato MARC21 Bibliográfico é dada pelas AACR2r.

Sintaxe e semântica em registros bibliográficos e o Formato MARC21

- 100 1# \$aSantos, Plácida Leopoldina Ventura Amorim da Costa
- 245 10 \$aCatalogação \$btrajetória para um código internacional
- 260 ## \$aNiterói \$bIntertexto \$c2009

Relação semântica:

- entre a TAG, os indicadores 1 e 2, os delimitadores e os códigos de subcampo e o conteúdo do campo, dados pelo Formato MARC21 Bibliográfico;
- entre os campos do registro bibliográfico (representante) em Formato MARC21 Bibliográfico e o representado.

Como se dá a conversão pela sintaxe e pela semântica?

- Utilizando-se de técnicas advindas da área de inteligência artificial
- Redes Bayesianas
- Representação de linguagem natural

Resultados esperados e contribuições

- Metodologia a ser implementada em ferramenta computacional, com uso de *software* livre, de código aberto
- Aspectos teórico-metodológicos da representação bibliográfica, com aspectos sintáticos e semânticos dos objetos representados.

Obrigada!

zzafalon@gmail.com

placidasantos@gmail.com