



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

Faculdade de Filosofia e Ciências
Departamento de Ciência da Informação
Campus de Marília

ADRIANA NASCIMENTO FLAMINO

**MARC21 e XML como ferramentas para a consolidação
da catalogação cooperativa automatizada:
uma revisão de literatura**

Marília – SP
2003

ADRIANA NASCIMENTO FLAMINO

**MARC21 e XML como ferramentas para a consolidação
da catalogação cooperativa automatizada:
uma revisão de literatura**

Trabalho de Conclusão de Curso apresentado ao Conselho de Curso de Biblioteconomia da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia.

Linha de pesquisa: Tecnologias em Informação.

Orientadora: Dr^a. Plácida Leopoldina Ventura Amorim da Costa Santos.

Bolsa de Iniciação Científica: Fundação de Amparo à Pesquisa do Estado de São Paulo – **FAPESP**.

Marília – SP

2003

F579m Flamino, Adriana Nascimento
MARC21 e XML como ferramentas para a consolidação da
catalogação cooperativa automatizada : uma revisão de literatura /
Adriana Nascimento Flamino. – Marília : UNESP, 2003
142 f : il. ; 30 cm

Trabalho de Conclusão de Curso - TCC (Graduação) –
Faculdade de Filosofia e Ciências – Universidade Estadual
Paulista, Marília, 2003.

1 Catalogação cooperativa automatizada. 2 MARC21. 3 XML.
4 Metadados. I. Autor II. Título.



CDD 025.4

ADRIANA NASCIMENTO FLAMINO

**MARC21 e XML como ferramentas para a consolidação
da catalogação cooperativa automatizada:
uma revisão de literatura**

Trabalho de Conclusão de Curso apresentado ao Conselho de Curso de Biblioteconomia da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia.

Linha de pesquisa: Tecnologias em Informação.

Bolsa de Iniciação Científica: Fundação de Amparo à Pesquisa do Estado de São Paulo – **FAPESP**.

Data de defesa : 1º / 12 / 2003

BANCA EXAMINADORA:

Nome: Plácida Leopoldina Ventura Amorim da Costa Santos

Titulação: Prof^ª. do Departamento de Ciência da Informação, Fac. de Filosofia e Ciências UNESP – Campus de Marília.

Nome: Silvana Aparecida Borsetti Gregório Vidotti

Titulação: Prof^ª. do Departamento de Ciência da Informação, Fac. de Filosofia e Ciências UNESP – Campus de Marília.

Nome: José Eduardo Santarém Segundo

Titulação: Mestrando em Ciência da Informação UNESP – Campus de Marília.

Local: Universidade Estadual Paulista -UNESP
Faculdade de Filosofia e Ciências
Campus de Marília

DEDICATÓRIA

Este trabalho é para mim a realização de um sonho e o resultado de muito esforço, perseverança e superação de obstáculos. Eu o dedico com muito amor e carinho a pessoas muito especiais em minha vida:

À Deus, pelo poder da vida e da inteligência humana;

Aos meus pais Maria Merseis Flamino, Carlos Gomes Flamino e José Magalhães Guedes, pelo incentivo aos estudos, pelo amor e compreensão;

Ao meu avô e também pai Abílio Francisco do Nascimento pelo exemplo de vida;

Aos meus irmãos Jair, Joel (in memória), Joelma, Rosilda, Roseane, Alcemar e Adenice pelo amor, pela paciência, amizade, e pela rica e maravilhosa experiência de ser irmã de todos vocês;

Aos meus sobrinhos Laura, Rafael, Isadora e Vitor, por vocês existirem e por proporcionar tantas alegrias a todos da família;

Aos meus priminhos Marcos Vinícius e Joyce, que continuem sempre assim, com muita dedicação aos estudos e a leitura;

Aos meus tios e padrinhos Luiz Ferreira do Nascimento, Enedite Nascimento Carneiro, Antônio Batista do Nascimento, e ao meu primo Ivanildo Francisco Carneiro, pelo carinho, compreensão e pela ajuda moral e financeira. Sem esse gesto maravilhoso de vocês seria quase que impossível a minha formação superior;

Aos meus primos Anderson, Andréia e Ariane, pela força, amizade, companheirismo, compreensão e paciência, principalmente quando fazíamos o cursinho CLQ-OBJETIVO juntos;

Aos meus futuros sogros João e Edite e ao meu futuro cunhado Ricardo pelo apoio, carinho, amizade e compreensão sempre.

À você, meu amor, Rodrigo Moreira Garcia, desde que eu o conheci há cinco anos atrás sempre esteve ao meu lado, sendo meu amigo, companheiro, namorado, hoje noivo e se Deus quiser e permitir, meu futuro marido. Que este trabalho seja para você um exemplo de luta e de conquista. E tenha sempre em mente a mensagem de Anais Ninn: “A vida se contrai e se expande proporcionalmente à coragem do indivíduo”.

Ao meu bebê, que ainda está no ventre, fruto do grande amor entre mim e o pai dele, Rodrigo Moreira Garcia.

AGRADECIMENTOS

Agradeço a todos que tornaram possível a realização deste trabalho e, em especial:

À Fundação de Amparo a Pesquisa do Estado de São Paulo – **FAPESP**, pelo apoio financeiro para o desenvolvimento desta pesquisa de Iniciação Científica;

À professora Plácida Leopoldina Ventura Amorim da Costa Santos, pelo apoio e estímulo e pela orientação objetiva;

À professora Silvana Aparecida Borseti Gregório Vidotti pelo estímulo e colaboração;

À professora e coordenadora geral de Bibliotecas Mariângela Spotti Fujita por oferecer a oportunidade de estagiar no processo de Conversão Retrospectiva da rede UNESP de Bibliotecas;

Às bibliotecárias da Coordenadoria Geral de Bibliotecas – CGB, Cássia Adriana de Sant’ Ana Gatti, Dilnei Fátima Fogolin, Maria Constância Martinhão Souto e Maria José Stefani Buttarello pelo apoio, carinho, compreensão e dedicação durante todo o meu estágio no Laboratório de Tecnologias Informacionais – LTI.

À professora e Consultora Sylvia Helena Morales Horiguela de Moraes (Escritório de Pesquisa da UNESP) pela pertinente orientação para a reformulação do meu projeto de pesquisa para o pedido de reconsideração à FAPESP, e a Marizinha (Kimiko Marisa Takahashi), pela atenção, disposição e carinho para comigo.

Ao Professor João Batista Ernesto de Moraes, pela amizade sincera e pela valiosa correção ortográfica e gramatical do meu projeto de pesquisa;

Ao José Eduardo Santarém Segundo, por fazer parte de minha banca examinadora e pelas pertinentes sugestões dadas ao meu trabalho de conclusão de curso.

Ao José Carlos Abbud Grácio (Zeca do STI) pela oportunidade de estagiar no Laboratório de Informática do campus e assim aperfeiçoar os meus conhecimentos nesta área, e pela sua compreensão, atenção e paciência durante todo o meu estágio.

À Marlene Gomes de Carvalho (Secretária do Diretor da Faculdade), por ser amiga nos momentos em que eu mais precisei.

À Vânia M. S. Reis Fantin (Diretora da Biblioteca do campus de Marília), Hiromi, Célia, Luzinete e Tina (Bibliotecárias), Zilda, Sônia, Lair, Ilma, Akiko, Neuza e Telma (funcionários) pelo apoio, colaboração, amizade e compreensão sempre.

À Stefanie Miller, bibliotecária da OCLC – Latin América and Caribbean, pelo material enviado a mim e por sempre responder as minhas dúvidas.

Aos Professores de Inglês da escola FISK de Marília, Vanessa, Renato, Luiza, Lilian e a Diretora Silvia pela forma alegre e muito didática de conduzir o ensino de inglês, pelas dicas, atenção e amizade.

Ao Sr. Adauto Justino Correa e família, pelo apoio, colaboração e amizade.

Às irmãs Rachel e Roberta (Pós-Graduação) que sempre foram tão atenciosas e simpáticas comigo sempre que precisei delas.

À todos os colegas de classe (XXIV turma de Biblioteconomia) pela rica experiência de vida.

“Estou sempre bem disposto a fazer bom juízo de um homem que tem o hábito da leitura. Ler é manter sempre vivas e despertadas as nobres faculdades do espírito, dando-lhes por alimento novas emoções, novas idéias e novos conhecimentos. Ler é multiplicar e enriquecer a vida interior.

Ler é sobretudo associar-se à existência de seus semelhantes, fazer ato de união e de fraternidade com os homens. O que lê, embora se encontre confinado em uma aldeia, vive o momento universal e pode dizer como o homem de terêncio: nada que é humano lhe é indiferente.

A leitura fecunda o coração, dando intensidade, calor e expansão aos sentimentos. Os egoístas, em geral, não se dão à leitura, porque vivem absorvidos na árida contemplação dos seus interesses pessoais. Não sentem necessidade de saírem de si mesmos e aproximarem-se dos seus semelhantes.

Os prazeres sociais encantam o homem; mas nem sempre, vêm a seu encontro nem dependem da sua vontade. Entretanto, os prazeres que proporciona a leitura são de todo o tempo e de qualquer lugar. E são os únicos que podemos renovar à nossa vontade”.

Nicolás Avellaneda.

Político e homem de letras argentino do século XIX.

É através da leitura e de seu entendimento que alcançamos o conhecimento, e através deste diminuimos as incertezas e realizamos as nossas tomadas de decisões mais acertadamente. As leituras foram as fontes de inspiração para a realização do meu projeto e conseqüentemente desta pesquisa. Como sempre diz a minha orientadora e professora Plácida L.V.A da Costa Santos, “sem leitura não há inspiração”.

FLAMINO, A. N. **MARC21 e XML como ferramentas para a consolidação da catalogação cooperativa**: uma revisão de literatura. 2003. 142 f. Trabalho de Conclusão de Curso – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília. 2003.

RESUMO

A cooperação e o compartilhamento de informações catalográficas e bibliográficas em meio automatizado, apenas se realizou com a criação e adoção do formato de Intercâmbio MARC21. Mas devido aos avanços das tecnologias de informação e comunicação, do crescente uso da Internet e das bases e bancos de dados, houve a necessidade da criação e desenvolvimento de ferramentas que otimizem cada vez mais as atividades de organização, recuperação e intercâmbio de informações. A XML é um desses desenvolvimentos que têm como propósito facilitar o gerenciamento, armazenamento e transmissão de dados via Internet. Diante disso, foi proposto através de uma revisão de literatura, analisar o Formato de Intercâmbio MARC21 e a Linguagem de Marcação XML como ferramentas para a consolidação da Catalogação Cooperativa Automatizada, suas diferenças de flexibilidades de armazenamento, organização, recuperação e intercâmbio de dados via Internet. Esta pesquisa possibilitou a divulgação à comunidade bibliotecária, através de uma revisão de literatura, do que vem sendo discutido internacionalmente sobre MARC21 e XML.

Palavras chave: Catalogação Cooperativa Automatizada; MARC21; Metadados; Tecnologias em Informação; XML.

FLAMINO, A. N. **MARC21 e XML como ferramentas para a consolidação da catalogação cooperativa**: uma revisão de literatura. 2003. 142 f. Trabalho de Conclusão de Curso - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília. 2003.

ABSTRACT

The cooperation and the sharing of information cataloguing and bibliographical in environment automated, this was only possible with the creation and adoption of interchange format MARC21. But due to the progresses of the technologies of information and communication, of the crescent use of Internet and of the databases and databanks, there were the need of the creation and development of tools that optimize the organization activities, retrieval and interchange of information. XML is one of those developments that have as purpose to facilitate the management, storage and transmission of data through Internet. Before that, it was proposed through a literature revision, to analyze Interchange Format MARC21 and Markup Language XML as tools for the consolidation of the Automated Cooperative Cataloguing, your differences of storage flexibilities, organization, retrieval and interchange of data through Internet. This research made possible the divulgation to the community librarian, through a literature revision, that has been discussed internationally on MARC21 and XML.

Keywords: Automated Cooperative Cataloguing; MARC21; Metadata; Technologies in information; XML.

LISTA DE ILUSTRAÇÕES

FIGURA 1:	Representação do item documentário	26
FIGURA 2:	Campo 245	41
FIGURA 3:	Campo 260	41
FIGURA 4:	Formato de Entrada	44
FIGURA 5:	Formato de Intercâmbio	46
FIGURA 6:	Cópia da figura 5	46
FIGURA 7:	Etiquetas MARC21 organizadas em forma de tabela	47
FIGURA 8:	Exibição com etiqueta	48
FIGURA 9:	Exibição formatada	49
FIGURA 10:	Campo 245 – indicador 1	49
FIGURA 11:	Campo 245 – indicador 1 igual a 0	50
FIGURA 12:	Registro com entradas corretas	51
FIGURA 13:	Registro aproveitável	52
FIGURA 14:	Registro sendo modificado	52
FIGURA 15:	Novo número de controle adicionado a copia do registro	53
FIGURA 16:	Registro alterado	53
FIGURA 17:	Registro sendo confirmado	54
FIGURA 18:	Registro salvo corretamente no banco de dados	54
FIGURA 19:	Etiquetas relacionadas	55
FIGURA 20:	Tabela da etiqueta 043	56
FIGURA 21:	Tabela da etiqueta 045	56
FIGURA 22:	Registro utilizando o campo 856	58
FIGURA 23:	Exemplos de documentos XML	84
FIGURA 24:	Exemplo de um documento em HTML	85
FIGURA 25:	Documentos em HTML	86
FIGURA 26:	Documento em XML	86
FIGURA 27:	Tela do editor XmlShell	91
FIGURA 28:	Exemplo de uma DTD interna	93
FIGURA 29:	Exemplo de uso de uma DTD externa	94
FIGURA 30:	Exemplo de um arquivo DTD	94
FIGURA 31:	Vários estilos de apresentação à partir de um mesmo documento XML	99
FIGURA 32:	Proposed namespace	104
FIGURA 33:	Áreas de aplicações XML	107
FIGURA 34:	Categorias de organizações envolvidas em aplicações XML	107
FIGURA 35:	Filosofia da XML	115
FIGURA 36:	Esquema MARC e XML	117
FIGURA 37:	O ciclo do aprisionamento	119

LISTA DE SIGLAS UTILIZADAS

AACR	Anglo-American Cataloguing Rules
ANSI	American National Standards Institute
CBU	Controle Bibliográfico Universal
CGB	Coordenadoria Geral de Bibliotecas
CSS	Cascading Style Sheets
DTD	Document Type Definition
FGV	Fundação Getúlio Vargas
GML	Generalized Markup Language
HTML	HyperText Markup Language
IFLA	Federação Internacional de Associações e Instituições Bibliotecárias
IMARC	International MARC
ISO	International Standardization Organization
LC	Library of Congress
LSP	Linked System Project
LTI	Laboratório de Tecnologias Informacionais
MARC	Machine Readable Cataloging
NISO	National Information Standards Organization
OCLC	Online Computer Library Center
RECON	Retrospective Conversion
RTF	Rich Text Format
SGML	Standard Generalized Markup Language
TCP/IP	Transmission Control Protocol / Internet Protocol
TEI	Text Encoding Initiative
UNESP	Universidade Estadual Paulista
UNIMARC	Universal MARC
W3C	World Wide Web Consortium
WLN	Western Library Network
XLink	XML Linking Language
XML	eXtensible Markup Language
XPath	XML Path Language
XPointer	XML Pointer Language
XSL	eXtensible Stylesheet Language
XSLFO	XSL Formating Object
XSLT	XSL Transformation

SUMÁRIO

1	INTRODUÇÃO	12
2	CATALOGAÇÃO COOPERATIVA AUTOMATIZADA	18
2.1	Histórico	21
2.2	Propósitos da Catalogação Cooperativa	27
2.3	Vantagens e desvantagens da Catalogação Cooperativa	27
2.4	Catalogação Cooperativa Automatizada	28
2.5	Bases e Bancos de dados em redes cooperativas	30
3	FORMATO DE INTERCÂMBIO MARC21	35
3.1	Histórico	36
3.2	Por que a necessidade de um único Formato MARC padrão?	38
3.3	Estrutura do registro MARC21	39
3.4	Facetas do Formato MARC21	43
3.5	Comunicação entre os Designadores de Conteúdos	49
3.6	Relações existentes entre etiquetas	55
3.7	Campo 856	57
3.8	Normas adotadas pelo MARC21	58
3.8.1	AACR2	59
3.8.2	Norma ISO 2709	60
3.8.3	Norma ANSI Z39.2	62
3.8.4	Protocolo de comunicação Z39.50	63
3.9	MARC21: Uma ferramenta , uma estrutura e um padrão para metadados	66
4	LINGUAGEM DE MARCAÇÃO XML	74
4.1	O que é XML?	75
4.2	Origem da XML	79
4.3	Vantagens da XML	81
4.4	Sintaxe básica da XML	83
4.4.1	Ferramentas para a criação de documentos XML	90
4.5	Tecnologias associadas à XML	91
4.5.1	DTD	92
4.5.2	XML Schema	96
4.5.3	Folhas de estilos CSS e XSL	98
4.5.4	XLink, XPointer e XPath	100
4.5.5	XML Namespaces	103
4.6	Uso da XML	105
5	ESTRUTURAS DE MARCAÇÃO MARC21 e XML	109
6	CONSIDERAÇÕES FINAIS	125
	REFERÊNCIAS	131
	BIBLIOGRAFIA CONSULTADA	136

CAPÍTULO 1

INTRODUÇÃO

Cooperação é a palavra-chave para todas as áreas no século XXI

A presente pesquisa de iniciação científica, intitulada **MARC21 e XML como ferramentas para a consolidação da Catalogação Cooperativa Automatizada: uma revisão de literatura**, inserida na linha de pesquisa Tecnologias em Informação, do Departamento de Ciência da Informação do Campus de Marília–SP, financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo – **FAPESP**, é apresentada como Trabalho de Conclusão de Curso, um dos requisitos para a obtenção do grau de Bacharel em Biblioteconomia, da Faculdade de Filosofia e Ciências da UNESP – Campus de Marília.

A Catalogação Cooperativa Automatizada, que hoje nos possibilita a cooperação e o compartilhamento de informações catalográficas e bibliográficas entre diversas instituições no ambiente automatizado somente se tornou possível com a criação e adoção do **Machine Readable Cataloging - MARC**, um Formato de Intercâmbio desenvolvido e testado pela **Library of Congress - LC** (Biblioteca do Congresso Norte Americano) na década de 1960.

Devido aos avanços das tecnologias em informação, do crescente uso da Internet e dos bancos de dados, houve a necessidade da criação de ferramentas que otimizem cada vez mais as atividades de organização, recuperação e intercâmbio de informações. Segundo Marcondes, (2001, p. 61) “de nada adianta a informação existir, se quem dela necessita não sabe da sua existência, ou se ela não puder ser encontrada”.

Carvalho já em 1990 alertava que,

A ciência da atualidade está produzindo uma quantidade de dados que vai atingir as proporções de um dilúvio, milhões de vezes superior à capacidade humana de manipulá-los ou estudá-los. Os pesquisadores produziram e armazenaram em máquina, nos últimos 20 anos, um volume de informações equivalente a tudo que foi descoberto nos 20 mil anos de história da civilização do homem (CARVALHO, 1990, p. 24).

Desta forma, novos formatos de intercâmbio e novas linguagens surgem. A **eXtensible Markup Language - XML** é um desses desenvolvimentos que tem como propósito facilitar o gerenciamento, armazenamento e transmissão de dados via Internet.

Portanto, para uma Unidade de Informação tornar-se útil é primordial que ela esteja conectada ao mundo globalizado, para assim recuperar, armazenar, representar e disseminar as informações atualizadas e necessárias ao usuário, independente de seu suporte informacional. Sendo assim, as ferramentas das novas tecnologias são as portas de entrada para essa realidade.

O interesse para a realização dessa pesquisa originou-se a partir da disciplina Catalogação Automatizada ministrada pela Professora Plácida Leopoldina Ventura Amorim da Costa Santos no ano de 2001, concretizando-se o interesse por este estudo na realização do estágio no processo de Conversão Retrospectiva (RECON), realizado no **Laboratório de Tecnologias Informacionais - LTI**, em conjunto com a **Coordenadoria Geral de Bibliotecas CGB** - da Universidade Estadual Paulista – UNESP - Faculdade de Filosofia e Ciências do Campus de Marília. Como parte das minhas atividades, participei da captação de registros bibliográficos em bases servidoras nacionais e internacionais e da importação dos mesmos para o banco de dados bibliográficos ATHENA da Rede UNESP.

O banco de dados reúne os registros bibliográficos de todas as bibliotecas da Rede UNESP, formando um catálogo coletivo pelo qual qualquer uma das unidades da Rede pode consultar e captar registros bibliográficos, além de contribuir com sua catalogação original.

Durante o estágio, pude perceber a importância de analisar o Formato de Intercâmbio MARC21 e a emergente Linguagem de Marcação Extendida - XML como ferramentas para a Catalogação Cooperativa, uma vez que, nossos instrumentos de trabalho eram Anglo-American Cataloguing Rules – 2ª edição - AACR2 e Machine Readable Cataloging - MARC21 para a efetivação da Catalogação Cooperativa Automatizada na UNESP.

A revisão de literatura é uma das metodologias de que os cientistas necessitam e se utilizam hoje para identificar, conhecer e acompanhar o desenvolvimento de pesquisas em suas áreas de atuação, pois a gama de informações existente hoje e a sua dispersão física dificultam cada vez mais este acompanhamento, tanto pela impossibilidade de se acessar tudo o que existe e controlar toda a bibliografia, como também pela limitação humana de absorver toda a informação relevante e disponível. As autoras Noronha e Ferreira (2000, p. 197) comentam com muita pertinência:

[...] os trabalhos de revisão, contribuem para o desenvolvimento do conhecimento científico, principalmente porque comparam informação de fontes diferentes; compactam o conhecimento existente; identificam especializações emergentes; direcionam pesquisas para novas áreas; notificam os pesquisadores periodicamente sobre a literatura publicada em dada época; promovem um serviço de alerta para campos correlatos; dão suporte à busca bibliográfica, constituindo-se em fonte inicial básica para a elaboração de projetos de pesquisa; auxiliam indiretamente o ensino, dando suporte a trabalhos acadêmicos e, finalmente, oferecem um feedback através da avaliação do trabalho publicado.

Pode-se perceber a necessidade de estudos como esse na área da Biblioteconomia, inclusive para que o profissional bibliotecário tenha mais conhecimento, domínio e utilize as atuais ferramentas de trabalho, os computadores, formatos de intercâmbio bibliográfico e

catalográfico, linguagens de marcação, protocolos e outros procedimentos envolvidos no processo da Catalogação Cooperativa Automatizada, do intercâmbio e do compartilhamento de dados bibliográficos e catalográficos.

Portanto, trata-se de um tema atual e necessário para a área da Biblioteconomia no que diz respeito à avaliação das formas de representação para o tratamento adequado das informações registradas.

O problema levantado e que motivou esta pesquisa foi o interesse em saber se, o **Formato de Intercâmbio MARC21 é imprescindível para a catalogação cooperativa automatizada**, uma vez que, o constante desenvolvimento de tecnologias em informação, juntamente com o crescente uso da Internet e dos bancos de dados, trazem a necessidade de se gerar novas ferramentas que otimizem cada vez mais as atividades de organização, recuperação e intercâmbio de informações.

A hipótese levantada para estudo foi: **A utilização do Formato MARC21 efetivaria uma catalogação cooperativa universal.**

Escolheu-se como objetivo geral ou final desta pesquisa, divulgar à comunidade bibliotecária através de uma revisão de literatura, o que está sendo discutido internacionalmente sobre MARC21 e XML. E como objetivos específicos, também, através de uma revisão de literatura, comparar o Formato de Intercâmbio Machine Readable Cataloging - **MARC21** e a eXtensible Markup Language - **XML** como ferramentas para a consolidação da Catalogação Cooperativa Automatizada, identificando as suas diferenças de flexibilidades de armazenamento, organização, representação, recuperação, e intercâmbio de dados via Internet.

Para realizar o presente estudo, procedeu-se da seguinte maneira:

Por se tratar de uma revisão de literatura, a mesma teve seu propósito, abrangência, função e tipo de análise desenvolvida.

Quanto ao seu propósito, este foi de base, ou seja, “servir de apoio, suporte para a comprovação ou não das hipóteses e idéias na pesquisa científica” (NORONHA; FERREIRA, 2000, p. 193). A sua abrangência foi temática, pois todo o trabalho de revisão foi calcado em um tema específico. Quanto a sua função, esta foi de atualização, pois foram coletadas e analisadas literaturas publicadas recentemente sobre o tema coberto. E o tipo de análise desenvolvida sobre as literaturas selecionadas foi bibliográfica, servindo como subsídio para comparação das diferentes abordagens dadas nos diversos trabalhos (NORONHA; FERREIRA, 2000). Portanto, a natureza desta investigação foi teórica, de nível descritivo e exploratório, e o tipo de pesquisa bibliográfica.

Para isso utilizou-se de fontes bibliográficas primárias, secundárias e terciárias, para levantamento bibliográfico.

Compreende-se por fontes bibliográficas primárias, obras produzidas com a interferência direta do autor da pesquisa como, livros, teses, TCCs, relatórios técnicos, normas técnicas, artigos científicos e outros. Como fontes secundárias, documentos produzidos a partir de documentos primários e são representadas por exemplo pelas enciclopédias, tratados, manuais, dicionários e outros. Como fontes terciárias documentos que tem por função guiar o usuário para as fontes primárias e secundárias como, bibliografias, catálogos coletivos, guias de literatura, diretórios, índices e outros (MUELLER, 2000).

O levantamento bibliográfico dos documentos primários, secundários e terciários para análise foi feito nas seguintes fontes bibliográficas da área de Ciência da Informação:

- Livros, periódicos, anais e relatórios ;
- Bases de dados nacionais e internacionais, textuais e referenciais (Web of Science, WebSPIRS, Current Contents, Probe, Scielo, Cruesp, Emerald dentre outras).
- Internet.

A pesquisa foi realizada através de leituras, análises e fichamentos dos documentos selecionados, nos idiomas inglês, espanhol e português. Utilizando-se de imagens de registros no formato MARC21 e no formato XML para ilustrações.

Com o intuito de alcançar os objetivos desta pesquisa, estruturou-se o trabalho em 6 capítulos, incluindo a Introdução que aborda a origem e justificativa da pesquisa, seu problema, hipótese, objetivos e metodologia.

Os seguintes capítulos estão assim organizados:

O capítulo 2: Catalogação Cooperativa Automatizada, apresenta seu histórico, conceitos, a filosofia cooperativa no contexto da Internet e dos bancos de dados, seu propósito, vantagens e desvantagens.

O Capítulo 3: Formato de Intercâmbio MARC21, apresenta seu histórico, a origem do Formato, conceitos, seus campos, sub-campos e indicadores, utiliza-se de figuras para ilustrar exemplos, aborda-se as normas ISO 2709, ANSI Z39.2 e o Protocolo de intercâmbio Z39.50, assim como as suas características e estrutura para metadados.

O Capítulo 4: eXtensible Markup Language - XML, apresenta o que é, a sua origem, suas vantagens, sua sintaxe básica, as ferramentas utilizadas na criação de documentos XML, abordando a SGML, as limitações da HTML e o poder da XML, suas tecnologias associadas como, DTD, XML Schema, Folhas de estilos CSS e XSL, XLink, XPointer e XPath , XML Namespaces e o uso da XML.

O Capítulo 5: MARC21 e XML, faz uma comparação entre as duas ferramentas, abordando as diferenças e semelhanças, suas vantagens e desvantagens. Utiliza-se de figuras para ilustrar exemplos.

O Capítulo 6: Considerações finais, relata os resultados alcançados por esta pesquisa.

CAPÍTULO 2

CATALOGAÇÃO COOPERATIVA AUTOMATIZADA

A introdução dos computadores nas bibliotecas resultou em padronização, aumento da eficiência, cooperação e melhores serviços.

Rowley

A informação vem assumindo, cada vez mais, uma posição estratégica, tornando-se um fator indispensável à sobrevivência das organizações, das sociedades e dos indivíduos. E uma das áreas da Biblioteconomia que tem por objetivo a representação e recuperação desta informação tão essencial hoje, é a catalogação que aos poucos vem obtendo o reconhecimento e a valorização por parte da sociedade uma vez que este é um trabalho imprescindível na interação com a mesma, pois possibilita a interdisciplinaridade, a disseminação e recuperação da informação, de forma padronizada, rápida e objetiva.

Com a introdução das tecnologias de informática e dos sistemas de comunicação na Biblioteconomia, tornou-se possível a utilização de padrões computacionais para a emissão e recepção da informação, permitindo a transferência e o compartilhamento de recursos informacionais através do intercâmbio de dados de descrição bibliográfica entre diversas instituições.

O processo para se realizar a catalogação de um item informacional não é fácil, despense tempo, custos, paciência e treinamento para a utilização de códigos, guias e padrões internacionais para que se tenha um grau ótimo de qualidade na padronização e normalização. O uso de tecnologias de informação na área da Biblioteconomia vem ao encontro de suas necessidades, ou seja, vem cada vez mais facilitar a vida e o trabalho dos profissionais da informação.

Para Lehnus (1971 p.11),

A catalogação descritiva apresenta, em geral, dificuldades tanto para os estudantes como para os que se iniciam nas tarefas de representação descritiva, não somente devido ao trabalho peculiar de registrar acuradamente e em forma pré-estabelecida, a descrição dos documentos, mas também pela individualização de cada [item informacional].

Mesmo catalogadores experientes consideram certos itens difíceis para serem descritos, e a causa pode ser não somente a falta de dados necessários, mas também o problema para distinguir entre o supérfluo e o essencial.

Ambos, o experimentado e o aprendiz, necessitam conhecer o Código de Regras e ter guias para sua orientação, a fim de descrever os documentos correta e uniformemente.

Segundo Lehnus (1971) nada melhor para compreender totalmente o significado de uma regra como ver a sua aplicação.

A necessidade de padronização e normalização hoje se estende a todos os campos e setores da vida moderna, sendo essencial em instituições ou organizações que atuam diretamente com a informação. “Porém no campo da catalogação é, além do mais, uma consequência dos progressos realizados pela Biblioteconomia, e que estão intimamente relacionados com o advento da informática” (LEHNUS, 1971, p.8).

E, para o desenvolvimento e gerenciamento com qualidade de catálogos, bases e bancos de dados de registros catalográficos e bibliográficos e, finalmente, a cooperação e o intercâmbio entre unidades de informação, a chave para o sucesso é a padronização. A cooperação e o intercâmbio de informações tornou-se essencial em qualquer instituição, dependendo cada vez mais destes fatores para a sua sobrevivência.

Segundo Cunha (1984) algumas bibliotecas até iniciarem suas atividades de intercâmbio e cooperação entre elas, a maioria tentavam ser tão auto-suficientes quanto

possível. “Sempre foi um elemento de prestígio e status, se ter uma coleção mais completa que a de outras instituições similares. Todavia, tal política é completamente impraticável [hoje em dia]” (CUNHA, 1984, p.42).

O momento em que estamos inseridos nessa nova era da informação exige colaboração, cooperação, intercâmbio, consórcios, sinergia, trabalho em equipe e mudança de paradigmas.

“Atualmente, fala-se muito em mudanças de paradigma, sobretudo quando avalia o impacto da Internet nas organizações. Mas o que se entende por mudanças de paradigma?” (LEITÃO; ROSSI, 2002, p. 78). Compreende-se por paradigma, “o conjunto de valores e crenças que ditam o comportamento e as ações coletivas em uma determinada época” (GRAMIGNA, 1995, p. 10). Desta forma, “os paradigmas definem uma forma de ver o mundo e de como ele deve ser estudado” (BURREL, 1999 citado por PITASSI; LEITÃO, 2002, p. 78).

“Assim, mudar um paradigma envolve mudar os pressupostos do conhecimento que o determina e também as crenças, os valores e a ideologia a ele associados” (LEITÃO; ROSSI, 2000, citados por PITASSI; LEITÃO, 2002, p. 78).

2.1 Histórico

[...] a atividade organizada voltada para a satisfação da curiosidade acerca do mundo natural é uma das definições de ciência. E, como observou Aristóteles, “todos os homens possuem por natureza o anseio do conhecimento”. Essa necessidade engloba até mesmo nosso anelo de sistematizar o universo. Jesse H. Shera lembrou-nos que o “homem abomina o caos, assim como se diz que a natureza abomina o vazio”, e procuramos incessantemente impor um padrão ao que vemos. (GROGAN, 1995, p. 22)

A necessidade que o homem tem de ordenar e padronizar o mundo a sua volta, remonta a tempos antigos, para se ter uma idéia, uma das listas de livros mais antiga que se tem conhecimento data de 2000 antes de Cristo. A catalogação é uma das formas que ele encontrou para organizar e ao mesmo tempo recuperar o conhecimento por ele produzido.

A catalogação consiste em representar itens documentários de forma clara, sucinta e padronizada, individualizando-os e tornando-os acessíveis por várias formas de acesso para que o usuário possa recuperá-lo com rapidez e objetividade. A catalogação tem por objetivo a recuperação da informação, isto é, dizer onde está o documento.

Entende-se por item documentário qualquer objeto passível de ser catalogado como, livros, CDs, fotos, partituras, documentos eletrônicos, obras de arte e outros.

Os bibliotecários elaboram representações desses itens, de forma a simplificar a busca e recuperação. Essas representações abrangem tanto o aspecto físico dos itens como seu conteúdo. Através dessas representações criam-se diversos instrumentos como, bibliografias, catálogos, boletins de serviço de alertas, entre outros (MEY, 1995).

Representar tem o mesmo significado de descrever, ou seja, ser a imagem ou a reprodução de algo, fazer a descrição fiel de um objeto ou um item documentário, tornando-o único e presente novamente. Segundo Marcondes (2001, p. 64) “Etimologicamente, representar significa fazer alguma coisa presente novamente”.

“O fato de que o Homem pode manipular signos (símbolos que se referem a fenômenos naturais) na ausência dos próprios fenômenos constitui uma diferença vital nas características dos animais” (FOSKETT, 1980, p. 24).

A representação é uma atividade essencial e característica do ser humano. Para (MARCONDES, 2001, p. 63) “esta atividade mental e cognitiva foi mais útil ao homem em sua sobrevivência que seus atributos de força física. No amanhecer da humanidade, elas jogaram um papel fundamental na sobrevivência do homem”.

Em passado recente, a catalogação era vista como uma técnica de elaborar catálogos – idéia muito restritiva e simplista, ou como técnica que listava itens – mais restritiva ainda. O que diferencia a catalogação de um inventário, listagem de itens, é o fato de não apenas caracterizar os itens, individualizando-os, tornando-os únicos entre os demais, como também de reuni-los por suas semelhanças (MEY, 1995, p.5).

Com o aperfeiçoamento da catalogação com o passar dos anos juntamente com a sofisticação das tecnologias usadas, várias instituições sentiram a necessidade de elaborarem e seguirem padrões para a organização e padronização da mesma.

No século XVIII foi criado o Código Francês, o primeiro código de catalogação, no século XIX as bibliotecas da Prússia criaram as Instruções Prussianas. No início do século XX surge o Código da Vaticana, o **International Standard Bibliographic Description - ISBD**, ou seja, Descrição Bibliográfica Internacional Normalizada, e por último surgiu o **Anglo-American Cataloguing Rules - AACR**, ou seja, Código de Catalogação Anglo-Americano, hoje utilizado por vários países, inclusive pelo Brasil.

Apesar da diversidade de códigos da época, e de cada Instituição fazer a catalogação de seus itens, cada uma a sua maneira, o Bibliotecário Charles Coffin Jewett

funcionário da Smithsonian Institution, uma instituição fundada nos Estados Unidos da América em 1846 e que tem como objetivo fazer com que o conhecimento humano aumente e se difunda, visualizou a possibilidade de construir um catálogo coletivo, do qual fariam parte todas as bibliotecas públicas norte-americanas, acreditando que este seria o primeiro passo para a formação de um catálogo Universal. Na época uma idéia, hoje uma realidade, ainda não concretizada de fato, na forma do **Controle Bibliográfico Universal – CBU**, “cujo objetivo é reunir e tornar disponíveis, de maneira eficiente, os registros da produção bibliográfica de todos os países, concretizando uma rede internacional de informação” (CAMPELLO ; MAGALHÃES, 1997, p. 4).

A palavra catálogo se origina do grego: Kata (de acordo com, sob, em baixo ou parte) e logos (ordem, razão). Segundo Ferreira (1993, p. 108), catálogo significa: “relação ou lista metódica, em geral alfabética, de pessoas ou coisas”. De acordo com Mey (1995, p. 8) não podemos concordar com definições desse tipo, para ela os catálogos são, de fato, canais de comunicação e prefere defini-los da seguinte forma:

Catálogo é um canal de comunicação estruturado, que veicula mensagens contidas nos itens, e sobre os itens, de um ou vários acervos, apresentando-as sob a forma codificada e organizada, agrupadas por semelhanças, aos usuários desse(s) acervo(s).

Como os catálogos mais comuns antigamente eram em forma de livros, que tinham um alto custo de produção e não possibilitavam uma atualização freqüente, Jewett propôs que os registros bibliográficos fossem gravados em chapas de esterotipia individual. Isso facilitaria a utilização do catálogo sem que a ordem alfabética fosse comprometida. Segundo ele, esse processo facilitaria e baratearia a produção de um catálogo coletivo impresso em forma de livro, se o sistema fosse adotado por todas as bibliotecas norte-americanas (MERCADANTE, 1995).

Para que o almejado catálogo coletivo nacional de Jewett fosse concretizado, ele propôs duas ações que agilizariam a sua formação: a catalogação cooperativa como o processo de catalogação que facilitaria a cooperação entre as bibliotecas e a esteriotipia como processo de impressão.

Sendo assim, há muitos anos atrás, Jewett, um visionário além do seu tempo, lançou as bases da catalogação cooperativa e que hoje pode se tornar uma realidade:

uma biblioteca jamais deveria catalogar novamente um material que já foi catalogado por outra biblioteca; para cada material que chega à mesa do catalogador, é necessário saber antes se alguém, em algum outro lugar do país ou do mundo, já o catalogou; se o material já tiver sido catalogado, todos os esforços devem ser enviados para se ter acesso a essa catalogação, e aproveitá-la (BALBY, 1995. p.30).

Para que possamos compreender e conceituar melhor a Catalogação Cooperativa é necessário que antes saibamos os significados de algumas palavras como, cooperar, cooperação, cooperativismo, cooperado e cooperativa.

A palavra **COOPERAR** deriva etimologicamente da palavra latina "Cooperari", formada por "cum" (com) e "operari" (trabalhar), e significa agir simultânea ou coletivamente para um mesmo fim, ou seja, trabalhar em comum para o êxito de um mesmo propósito.

COOPERAÇÃO é o método de ação pelo qual indivíduos ou famílias com interesses comuns constituem um empreendimento. Neste, os direitos de todos são iguais e o resultado alcançado é repartido somente entre os integrantes, na proporção da participação societária nas atividades.

COOPERATIVISMO é uma doutrina, um sistema, um movimento ou simplesmente uma atitude ou disposição que considera as cooperativas como uma forma ideal de organização das atividades sócio-econômicas da humanidade.

COOPERADO é o produtor rural, o trabalhador urbano ou outro profissional, de qualquer atividade sócio econômica, que se associa para participar ativamente de uma cooperativa, cumprindo com os deveres e observando os seus direitos.

COOPERATIVA é uma associação de pessoas com interesses comuns, organizada economicamente e de forma democrática, com a participação livre de todos os que têm idênticas necessidades e interesses, com igualdade de deveres e direitos para a execução de quaisquer atividades, operações ou serviços. (Significados elaborados a partir de obras de referência como, dicionários e enciclopédias).

Depois de todas essas definições podemos conceituar **Catálogo Cooperativa** como, o trabalho de Catalogação realizado de forma cooperativa, ou seja, trabalho realizado coletivamente por pessoas ou instituições que tenham objetivos em comum para o êxito de um mesmo propósito. Sendo assim, cada pessoa ou unidade de informação tem os seus direitos ao mesmo tempo que tem deveres a cumprir. As unidades de informação inseridas neste sistema tem participação ativa na confecção dos registros catalográficos e bibliográficos.

Os registros catalográficos tem por objetivo a recuperação da informação, ou seja, dizer onde está o documento, diferentemente do registro bibliográfico que tem por objetivo disseminar a informação sobre o documento, ou seja, dizer que o item informacional existe. Observe as diferenças:

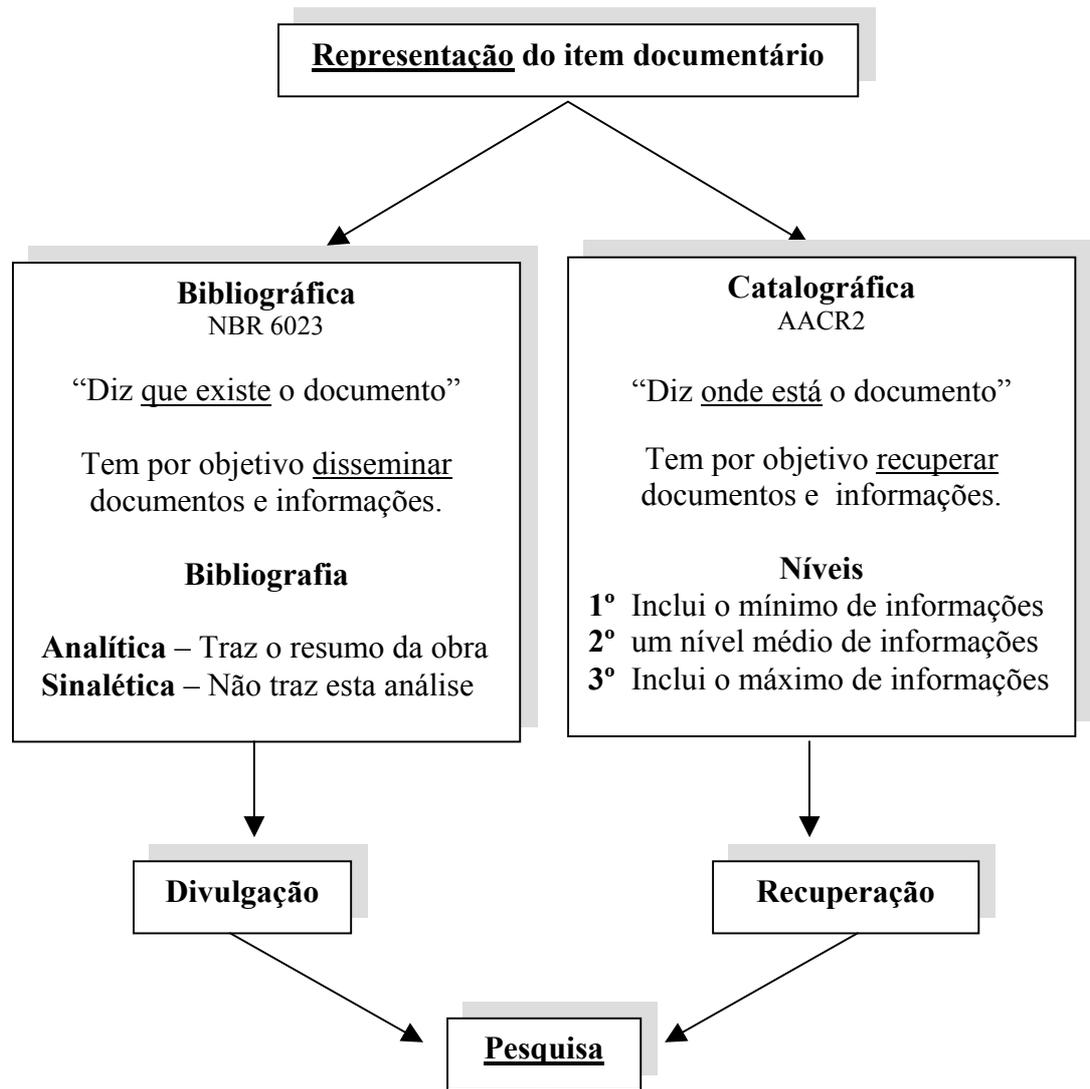


FIGURA 1: Representação do item documentário

A pesquisa (recuperação de informações através de estratégias de busca) é uma atividade especificamente do bibliotecário e que é prontamente reconhecida pelo usuário. No entanto, é necessário que primeiro seja feito o tratamento e a organização da informação, com instrumentos de representação como, a catalogação, a classificação e a indexação para que desta forma seja feito o seu correto armazenamento com o objetivo de possibilitar a recuperação da informação (SANTOS, 2002).

2.2 Propósitos da Catalogação Cooperativa

Os propósitos da Catalogação Cooperativa vão além de compartilhar e cooperar com registros catalográficos e bibliográficos entre instituições. São eles:

- Aperfeiçoar os sistemas de catalogação;
- Favorecer a padronização de processos catalográficos;
- Contribuir para melhorar as condições de pesquisa bibliográfica;
- Ampliar as comunicações entre as bibliotecas;
- Estabelecer o serviço de intercâmbio de catalogação;
- Reduzir custos e tarefas no processo de catalogação;
- Estabelecer e auxiliar a organização de catálogos coletivos;
- Promover a divulgação da coleção existente em cada unidade participante.

(BALBY; GROVER; MERCADANTE, 1995).

Segundo Sambaquy (1951, p. 39), “O que importa é que as bibliotecas compreendam que, para servir bem, não lhes é necessário conhecer somente que livros possuem, mas também onde se encontra o livro ou a informação que realmente está sendo desejada”.

2.3 Vantagens e desvantagens da Catalogação Cooperativa

As vantagens da Catalogação Cooperativa são inúmeras, dentre elas podemos citar:

- Compartilhamento de dados bibliográficos;
- Redução nos custos dos serviços de catalogação;
- O acesso à informação mais amplo e mais rápido;
- Uma maior especialização do pessoal bibliotecário resulta na divisão do trabalho que ocorre em um sistema cooperativo;
- Padronização dos dados descritivos;

- Intercâmbio bibliográfico de alcance internacional (GROVER; MERCADANTE, 1995).

Como todo empreendimento humano, a Catalogação Cooperativa tem as suas vantagens e desvantagens. Para Mercadante (1995, p. 25), “Parece que a grande barreira a se enfrentar em relação ao compartilhamento da catalogação diz respeito à aceitação da filosofia do trabalho cooperativo. Ao se vincular à uma rede obterá consistência e conseqüentemente se fortalecerá através de seus membros”.

Muitos bibliotecários ainda consideram natural ou pelo menos inevitável, fazer a catalogação original de um material, sem antes procurar saber se em algum lugar do país ou do mundo, alguém já o catalogou, poupando assim o seu tempo e esforço e ao mesmo tempo reduzindo o custo da catalogação em sua instituição.

A Catalogação Cooperativa nos proporciona mais vantagens do que desvantagens. As suas desvantagens são:

- Apresenta problemas com relação às barreiras lingüísticas devido à necessidade de traduzir e adaptar notas e cabeçalhos de assunto de acordo com a língua oficial usada pelo país em que forem utilizadas;
- Diversidade de códigos de catalogação utilizados, pois alguns países ainda adotam códigos como as Instruções Prussianas, Código da Vaticana, entre outros, que oferecem várias opções para decisão do cabeçalho principal, seja pessoal ou coletivo.

2.4 Catalogação Cooperativa Automatizada

A cooperação sempre existiu na história humana, como uma alternativa de sobrevivência ou uma solução para os momentos de crise. Depois da segunda guerra mundial, na década de 1940, houve uma produção espantosa de documentos, gerando um grande

volume de informação, que foi designado nessa época como “ a explosão da informação”, e, a mesma, era parcialmente inacessível aos estudantes, cientistas e outros usuários, até o advento do computador e conseqüentemente da Internet.

A introdução e aplicação das tecnologias de informática para o processamento da informação bibliográfica constituiu um dos eventos mais importantes para a Biblioteconomia nas últimas décadas. Atualmente bilhões de referências são armazenadas e tornadas disponíveis a cada ano, sob a forma digital.

No contexto histórico em que estamos vivendo, a era da Internet, a rede das redes e da informação, a ação cooperativa ganha enorme importância. “A Internet se parece cada vez mais com o sonho de Otlet (1934) de uma biblioteca universal” (OTLET, 1934 citado por MARCONDES, 2001, p.68). “E para que ela não evolua no sentido de um caos informacional, tem de ser trabalhada cooperativamente pelos profissionais de informação de todo o mundo” (MARCONDES, 2001, p. 68). Hartung (1996, p. 7) cita os valores do cooperativismo que são: “ ajuda mútua; responsabilidade; democracia; igualdade; equidade; solidariedade; honestidade; transparência; responsabilidade social e preocupação pelo seu semelhante”.

Para Souza (1999, p. 248),

A cooperação e o compartilhamento entre unidades de informação tem sido apresentados como a solução ideal para que possam atingir altos níveis de eficiência na busca e obtenção das informações necessárias a usuários que buscam subsídios para a realização de seus trabalhos, estudos, apoio ao ensino, ao desenvolvimento de pesquisas, à realização de negócios e outros.

No que concerne à catalogação, as vantagens da cooperação são inúmeras e os bibliotecários, pelo menos na sua maioria, já se conscientizaram de que é uma grande perda de tempo, energia e dinheiro, duplicar a catalogação dos mesmos documentos adquiridos por diversas unidades de informação no país.

A cooperação e o compartilhamento de informações catalográficas e bibliográficas entre diversas instituições no ambiente automatizado somente se tornou possível com o desenvolvimento e adoção do **Machine Readable Cataloguing - MARC**, um Formato de Intercâmbio desenvolvido e testado pela **Library of Congress - LC** (Biblioteca do Congresso Norte Americano) na década de 1960. Este formato é considerado uma ferramenta importante que tem como objetivo conseguir um grau ótimo de normalização, padronização da representação descritiva automatizada dos dados informacionais, compatibilidade e transferência de dados legíveis por computador, permitindo a cooperação e o compartilhamento de recursos e serviços entre diferentes unidades de informação.

No próximo capítulo estaremos abordando o Formato com maiores detalhes.

Tarapanoff (1996) citado por Silva (1999/2000, p. 5) considera que,

Os tipos mais conhecidos de unidades de informação são as bibliotecas, os arquivos públicos e os museus, estando inseridos neles os serviços de referência, documentação e informação. Acrescentam-se a eles, em consequência do desenvolvimento da informática, outros tipos de serviços capazes de fornecer informações a usuários que são as empresas que desenvolvem bases de dados específicas, destinadas ao atendimento de clientelas específicas.

2.5 Bases e Bancos de dados em redes cooperativas

As bases de dados significam um conjunto de informações apresentadas de forma organizada e padronizada, e com um objetivo específico, seja em ambiente digital ou analógico. As bases de dados são muito comuns no nosso dia-a-dia e, raramente, paramos

para pensar nisso. Exemplos de bases de dados que não seja do computador são as listas telefônicas, os dicionários, nossa agenda telefônica, as enciclopédias e outros.

De acordo com Lancaster, (1993, p. 305), podemos conceituar base de dados como:

uma coleção de itens sobre os quais podem ser realizadas buscas com a finalidade de revelar aqueles que tratam de um determinado assunto. A base de dados consiste em artefatos, como livros (o acervo de uma biblioteca é uma base de dados com certeza), ou registros que representam os artefatos, como por exemplo, registros bibliográficos constantes de páginas impressas, de fichas ou de meios eletrônicos.

Para Rowley (1994, p. 66) “Base de dados é uma coleção de registros similares entre si e que contém determinadas relações entre esses registros. Um sistema de base de dados pode abranger várias bases interligadas”. A mesma autora, em 2002, conceitua base de dados como “uma coleção geral e integrada de dados junto com a descrição deles, gerenciada de forma a atender a diferentes necessidades de seus usuários” (ROWLEY, 2002, p. 125).

De acordo com Cianconi (1987) o conceito de base de dados e banco de dados é tratado pela literatura especializada como sinônimos, causando grande confusão de conceitos. Porém, a autora aponta que, segundo a maioria da literatura na área de informação, pode-se conceituar base de dados como “um conjunto de dados interrelacionados, organizados de forma a permitir recuperação de informações”. E o termo banco de dados, segundo a autora, freqüentemente conceituado como sinônimo de base de dados pode ser visto como “um conjunto de bases de dados”.

Já os autores Guinchat e Menou (1994) consideram base de dados diferente de banco de dados, conceituando-os da seguinte forma: “uma base de dados é um conjunto organizado de referências bibliográficas de documentos que se encontram armazenados fisicamente em vários locais”. Segundo os autores, “Um usuário que busca uma informação

em uma base de dados bibliográfica recupera uma referência”. Na opinião dos autores “os bancos de dados tratam das informações factuais, numéricas ou textuais diretamente utilizáveis”. Ou seja, os bancos de dados, segundo os autores, proporcionam acesso direto às informações primárias.

Heemann (1997, p. 14) em sua dissertação de mestrado segue as conceituações de Cianconi (1987) e Rowley (1994) “que tratam as bases de dados como armazenadoras de registros e/ou dados visando sua recuperação”. Segundo a autora “essas abordagens são as mais utilizadas na Ciência da Informação, e pelos usuários em geral, para se referir a esse tipo de aplicativo”. Neste trabalho também serão consideradas as conceituações de Cianconi e Rowley.

As bases e bancos de dados em redes cooperativas são usadas para a cooperação e compartilhamento de registros catalográficos e bibliográficos entre si, para assim eliminar o contínuo acúmulo de materiais não catalogados e reduzir a quantidade de catalogações originais a serem feitas em suas instituições. Desta forma, dados bibliográficos vêm sendo adicionados às bases de dados cooperativas, transformando-as assim em reais e valiosos catálogos coletivos em linha, ou como são mais conhecidos: Catálogos on-line.

Os catálogos on-line têm a mesma estrutura que um catálogo manual, apenas encontram-se no formato digital e em redes, facilitando e multidimensionando o acesso do usuário. Os catálogos on-line geralmente encontram-se em formato MARC21, utilizando a estrutura de campos, subcampos e indicadores para descrição e o código de catalogação AACR2 para normalização e padronização dos dados. Os catálogos manuais são montados como o próprio nome já indica, manualmente, seguindo as regras do código de catalogação AACR2 para descrição, normalização e padronização dos dados.

Para Qin (2000), hoje os catálogos das bibliotecas estão longe de serem só uma ferramenta para suas próprias coleções, para uso somente dos visitantes das bibliotecas;

Estes se tornaram vários nós na rede, onde os usuários podem acessá-los de qualquer lugar do mundo através de um computador conectado à Internet.

Trabalhar em rede traz grandes desafios pessoais e profissionais, pois a evolução no domínio das técnicas de comunicação, o uso habilidoso e criativo das ferramentas tecnológicas, a revolução cultural, a internalização dos fundamentos, não podem ser processos apenas individuais, têm que ser coletivos, pois não se faz uma rede sozinho. Se há um espaço em que não se cresce sozinho é o das redes. E compartilhar [e cooperar] é a estratégia do crescimento conjunto (AMARAL, 2002, p. 1).

No Brasil temos a base de dados nacional BIBLIODATA, de responsabilidade da Fundação Getúlio Vargas (FGV), a qual utiliza-se do sistema de catalogação cooperativa, tendo a participação de várias Instituições brasileiras como cooperantes, sendo o seu catálogo coletivo on-line reconhecido internacionalmente.

No exterior as bases de dados⁴ mais importantes são: A Online Computer Library Center - OCLC e a Library of Congress – LC (Biblioteca do Congresso Norte Americano), entre outras. O catálogo coletivo on-line da OCLC oferece acesso a partir de pagamento, ou seja, a Instituição que tenha o interesse em compartilhar e/ou cooperar com registros catalográficos e bibliográficos com a mesma, precisará antes firmar contrato entre elas, e isso depende custos. Já o catálogo coletivo on-line da LC está disponível na Internet gratuitamente, no endereço <http://catalog.loc.gov>, para busca e compartilhamento de registros catalográficos e bibliográficos por toda e qualquer instituição que necessitar ou quiser.

As Instituições podem fazer parte das redes cooperativas de três formas:

- Compartilhando com registros catalográficos e bibliográficos de instituições;
- Gerando e distribuindo registros catalográficos e bibliográficos para instituições;

⁴ As instituições OCLC e LC consideram seus catálogos on-line como database, ou seja, base de dados.

- Cooperando com registros catalográficos e bibliográficos entre instituições.

Quando uma instituição compartilha um registro com outra, quer dizer que ela apenas capturou o registro dessa, para adicioná-lo em seu catálogo. Agora, quando uma instituição coopera registros com outra, quer dizer que ela disponibilizou os seus registros a outra instituição e portanto, tem o direito de capturar registros desta. Algumas instituições somente compartilham, outras só geram e distribuem registros, e outras cooperam, ou seja, tanto capturam os registros para si, como também dispõem dos seus para as outras instituições.

Neste contexto, a palavra **compartilhar** significa: ter ou tomar parte em. E **cooperar** significa: colaborar, auxiliar.

Desta forma, a semente filosófica da catalogação cooperativa lançada por Jewett, nos primórdios do século XIX, só agora está sendo relativamente aceita e internalizada não só pelos profissionais da informação como também de outras áreas como a medicina, a informática e outras, concretizando assim uma nova forma de organização, utilizando para isso a Internet e suas ferramentas.

No próximo capítulo estaremos abordando com maiores detalhes o formato MARC, a ferramenta que possibilitou a concretização da Catalogação Cooperativa em ambiente digital e em rede.

CAPÍTULO 3

FORMATO DE INTERCÂMBIO MARC21

Os computadores são incrivelmente rápidos, precisos e burros; os homens são incrivelmente lentos, imprecisos e brilhantes; juntos, seu poder ultrapassa os limites da imaginação.

Albert Einstein

Ouvimos diariamente e de forma natural as pessoas falarem coisas como, esse texto está em formato PDF, formato HTML, ou formato DOC, fiz o download de uma figura no formato JPEG, ou, salvei a figura que estava em formato JPG para formato GIF. Acreditamos porém que em nenhum momento essas pessoas pararam para pensar o que seria um formato ou qual seria seu conceito.

Formato, como a própria palavra nos sugere, é uma forma. Segundo Ferreira (1993, p. 258), “formato é a forma de apresentação de um valor; tipo determinado sob cujo modelo se faz algo; molde”.

Para Zocco (1993, p.5),

[...] um formato é um receptáculo ou estrutura na qual os dados que descrevem uma obra, como autor, título, assunto, edição, local de publicação etc., são inseridos. Esse receptáculo contém campos e são esses campos que determinam que cada um dos dados citados será inserido, ou seja, existe um campo para cada dado descritor da obra.

Desta forma, sabendo-se que o significado de intercâmbio é o mesmo que troca ou permuta, podemos definir de uma forma bem simples formato de intercâmbio catalográfico

e bibliográfico como, um molde ou forma para troca ou permuta de dados catalográficos e bibliográficos em ambientes digitais e/ou automatizados.

3.1 Histórico

O Formato de Intercâmbio **Machine Readable Cataloging - MARC** ou seja, um registro catalográfico legível por computador, foi criado a partir da necessidade de minimizar esforços, reduzir custos dos processos técnicos e compartilhar informações. Como sabemos, toda invenção ou criação humana nasce da necessidade de facilitar a vida cotidiana e beneficiar o próprio ser humano, e com o formato MARC não foi diferente.

Devido a explosão de publicações científicas após a segunda guerra mundial, buscava-se um meio de controlar, registrar e armazenar toda informação disponível. A década de 1960 surge com o desenvolvimento dos recursos computacionais, e a **Library of Congress - LC** desenvolveu o formato MARC durante 1965-1966. Isto conduziu a um projeto piloto, conhecido como MARC I, que investigou a praticabilidade de produzir dados de catálogos em formulários legíveis por computador.

Um trabalho similar estava em andamento no Reino Unido onde o Conselho de Bibliografia Nacional Britânica ajustou o projeto acima chamando-o de BNB MARC para examinar o uso de dados legíveis por computador na produção da Bibliografia Nacional Britânica impressa (BNB).

Em 1968 o projeto MARC II começou como um esforço Anglo-Americano para desenvolver um formato padrão de comunicação, ou seja, intercâmbio. Apesar da cooperação emergiram duas versões, UKMARC e USMARC.

O formato MARC torna-se base para outros formatos, a família cresce vertiginosamente e na década de 1970 contava-se mais de cinquenta formatos. Listaremos aqui alguns exemplos:

USMARC – Estados Unidos, **UKMARC** - Inglaterra, **MONOCLE** - França, **CANMARC** - Canadá, **MARC/BR** - Bélgica, **IBERMARC** - Espanha, **MARC/MÉXICO** – México, **SAMARC** – África do Sul, **ANB MARC** – Austrália, **MAB1** – Alemanha, **DN/MARC** – Dinamarca, **PICA** – Holanda, **NISSAT** – Índia, **MARCIS** – Israel, **ANNA MARC** – Itália, **JAPAN MARC** – Japão, **SWEMARC** – Suécia, **NORMARC** – Noruega, **CALCO** – Brasil, e muitos outros.

Com tantos formatos, inevitavelmente surgiram diferenças e dificuldades no intercâmbio de informações. Para solucionar o problema de incompatibilidade a Library of Congress cria um formato internacional do MARC, o UNIMARC (MARC Universal).

Este formato cobriu originalmente monografias e séries, mas foi estendido em 1987 para incluir materiais não livros. O formato foi mantido e desenvolvido sob a supervisão do Comitê permanente do UNIMARC.

Devido as razões econômicas da Biblioteca Britânica (UKMARC), da Biblioteca do Congresso Americano (USMARC) e da Biblioteca Nacional do Canadá (CAN/MARC) deu-se início ao movimento para a harmonização desses três formatos MARC buscando formas de reduzir os custos da catalogação por meio da harmonização entre eles, para uma simplificação e padronização da catalogação, e desta forma, uma biblioteca poderia utilizar os registros preparados por outras bibliotecas e assim reduzir custos nas instituições.

Depois de alguns estudos e tentativas, chegaram a conclusão que o formato UKMARC não poderia ser harmonizado completamente, em consequência das várias diferenças entre o UKMARC, o USMARC e o CAN/MARC. Alcançar a harmonização completa significaria que qualquer biblioteca e vendedores do sistema no Reino Unido teriam que fazer mudanças significativas ou os vendedores do sistema Norte Americano teriam que fazer mudanças caras no formato. Desta forma, concordaram em excluir o UKMARC da harmonização dos formatos.

Iniciou-se então em 1994 a harmonização dos dois formatos nacionais, o USMARC que foi desenvolvido na década de 1960 pela Library of Congress – LC (Biblioteca do Congresso Norte Americano), e o CAN/MARC desenvolvido em 1973 pela Biblioteca Nacional do Canadá, para controlar a sua catalogação legível por máquina.

Inicialmente foi utilizado o nome provisório de IMARC para o formato harmonizado, ou seja, MARC Internacional. Porém, durante a conferência da Federação Internacional de Associações e Instituições Bibliotecárias - IFLA em 1996, o nome IMARC causou uma reação muito negativa em vários países europeus, e inclusive da IFLA, pois o mesmo foi visto como uma competição com o formato UNIMARC o qual foi patrocinado pela IFLA. Certamente não foi a intenção do Canadá (CAN/MARC) competir com o UNIMARC. Assim, o IMARC estava descartado como um nome para os formatos harmonizados e depois de dois anos buscando possíveis nomes, em outubro de 1998, a Library of Congress e o Escritório de padrões MARC propuseram o nome MARC21 que significa MARC para o século 21. Este nome foi aceito, gerando assim um grande alívio a todos (STEWART, 1999).

MARC21 é o novo nome dos formatos harmonizados do CAN/MARC e do USMARC.

3.2 Por que a necessidade de um único Formato MARC padrão?

Cada Instituição poderia criar seus próprios métodos de organizar as informações bibliográficas, mas estaria deixando sua unidade de informação isolada, ilhada, limitando suas opções e criando mais trabalho para si mesmo. O uso do padrão MARC21 evita duplicação de trabalho e permite o melhor compartilhamento de recursos bibliográficos entre bibliotecas. A opção pelo uso do MARC21 permite que as bibliotecas adquiram dados catalográficos previsíveis e confiáveis. Se uma biblioteca ou centro de documentação decide utilizar um

sistema de automação desenvolvido localmente que não utilize registros MARC21, esta biblioteca ou centro de documentação não estará usufruindo da vantagem de se utilizar um padrão de amplo emprego no setor cujo propósito primário é promover a comunicação de informação (FURRIE, 2000).

3.3 Estrutura do registro MARC21

A estrutura de um registro MARC21 é composta de três elementos básicos. São eles: O **Líder**, o **Diretório** e os **Campos**.

O Líder - é o primeiro campo de um registro MARC, consiste dos primeiros 24 caracteres do registro e contém dados que fornecem informações para o processamento do registro. Cada posição tem um significado, mas muitas informações do líder são para uso do próprio sistema computacional.

Diretório - Imediatamente após o líder, vem um bloco de dados chamado diretório. Ele informa que etiquetas estão no registro e onde estão localizadas (dando o número do caractere onde cada campo começa). O Diretório é construído pelo computador a partir de um registro bibliográfico, baseado na informação de catalogação. Se qualquer informação catalográfica for alterada, o diretório pode ser reconstruído da mesma forma.

Campos – Cada registro bibliográfico é dividido logicamente em campos. Há um campo para autor, um campo para a informação do título e assim por diante. Esses campos são subdivididos em um ou mais “subcampos” e definidos por “indicadores”. Os nomes dos campos são representados por etiquetas de 3 dígitos. Apesar dos catálogos online exibirem os nomes dos campos como, autor, título etc., estes nomes são fornecidos pelo programa de computador que está sendo usado e não pelo registro MARC21.

Cada campo é associado a um número de 3 dígitos chamado “etiqueta”. Uma etiqueta identifica o campo, o tipo de dado que se segue. Embora a exibição destes dados por uma impressora ou pela tela do computador possa mostrar a etiqueta imediatamente seguida de indicadores, fazendo com que a etiqueta pareça ter 4 ou 5 dígitos, a etiqueta sempre se constitui dos primeiros três dígitos.

As etiquetas usadas com mais frequência são:

020 International Standard Book Number (ISBN)

100 Entrada principal pelo nome pessoal (Autor)

245 Informação de título (que inclui o título propriamente dito, outras informações sobre o título e menção de responsabilidade)

250 Edição

260 Informação de publicação, ou seja, imprensa (local, editora e ano)

300 Descrição física

440 Menção de série

500 Notas

650 Cabeçalho de assunto tópico

700 Entrada secundária de nome pessoal como, editor, tradutor, colaborador ou ilustrador.

Os exemplos dessas etiquetas estão relacionados a livros, mas a abrangência do formato MARC21 engloba não só livros como também materiais visuais, periódicos, materiais de arquivos e manuscritos, arquivos de computador, mapas, músicas e outros, e para cada tipo de material modificam-se algumas etiquetas para melhor descrever a especificidade de cada material.

Observe abaixo (FIG. 2) o exemplo de um campo. O número 245 é a etiqueta do campo que o define como o campo de título da obra.

código de subcampo. Os códigos de subcampos são letras minúsculas como pode ser visto na (FIG.3) : os subcampos representados pelas letras “a”, “b” e “c” e cada código de subcampo indica que tipo de dado o seguirá. No exemplo da (FIG.3) o código de subcampo “a” significa que o dado que se seguirá será o local de publicação, no subcampo “b” indica a casa publicadora e o subcampo “c” indica a data de publicação.

Os três tipos de designadores de conteúdo são símbolos abreviados que rotulam e explicam o registro bibliográfico. Designadores de Conteúdo é um termo abrangente usado para se referir às etiquetas, indicadores e códigos de subcampos (FURRIE, 2000).

As etiquetas do MARC21 estão divididas em centenas, as divisões básicas são:

- 0XX** Informação de controle, números, códigos
- 1XX** Entradas Principais
- 2XX** Títulos, edição, imprensa (em geral, o título, menção de responsabilidade, edição e informação de publicação)
- 3XX** Descrição física etc.
- 4XX** Série
- 5XX** Notas
- 6XX** Campos de assunto
- 7XX** Entradas secundárias
- 8XX** Entradas secundárias de série etc.
- 9XX** Livres para usos locais

A notação XX é freqüentemente usada para se referir a um grupo de etiquetas relacionadas. Por exemplo: 2XX refere-se a todas a etiquetas cujo primeiro dígito é “2” como, 210, 222, 240, 250 e assim por diante.

Os pontos de acesso, ou seja, entrada principal, entradas secundárias de assunto e outras entradas secundárias, são partes importantes de um registro bibliográfico, pois possibilitam ao usuário recuperar as informações contidas nos catálogos. A maioria dos pontos de acesso segundo Furrie (2000, p. 28) estão nos seguintes campos:

- 1XX** Entrada principal
- 4XX** Série
- 6XX** Cabeçalhos de assunto
- 7XX** Entradas secundárias
- 8XX** Entradas secundárias de série

Portanto, estes são os campos que estão sujeitos a controle de autoridade. Controle de Autoridade significa adotar uma forma autorizada ou estabelecida. É utilizado para isso uma lista ou base de dados com entradas de nome e assunto de cabeçalhos oficial (FURRIE, 2000). A UNESP por exemplo, utiliza a base de dados BIBLIODATA (Fundação Getúlio Vargas) como base oficial para entrada de assuntos e autoridades para pontos de acesso em seu Banco de Dados ATHENA.

3.4 Facetas do Formato MARC21

O Formato de Intercâmbio MARC21 se desdobra em três facetas, são elas: Formato de Entrada, Formato de Armazenamento Interno e Formato de Intercâmbio.

O Formato de Entrada é usado pelo catalogador na entrada dos dados. Para Balby (1995, p. 31) “um bom formato de entrada deve ser fácil de entender, ser próximo da forma de trabalho de quem coleta dados para o sistema, e possibilitar o preenchimento imediato, sem recurso a manuais separados de instruções.” O MARC21 parece seguir essa orientação. Veja a figura 4.

Lider	<u>LDR</u>	---	-----22-----4500
CPD[USE CTRL+F]	<u>001</u>	---	-----
Dados fixos	<u>008</u>	---	-----
ISBN	<u>020</u>	---	a
Fonte catal.	<u>040</u>	---	a
			b
Idioma [F8]	<u>041</u>	?	a
			b
Area geog. [F8]	<u>043</u>	---	a
Per. cron.[F2]	<u>045</u>	---	a
N. CDU	<u>080</u>	---	a
N. CDD	<u>082</u>	<u>04</u>	a
			z
Nome pessoal	<u>100</u>	?	a
			b
			c
			d
			e
			f
			g
Nome corpor.	<u>110</u>	?	a
			b
			c
			d
Nome evento	<u>111</u>	?	a
			c
			d
			n
Titulo	<u>245</u>	<u>1?</u>	a
			b
			c
Edição	<u>250</u>	---	a
			b
Imprensa	<u>260</u>	---	a
			b
			c
Desc.fisica	<u>300</u>	---	a
			b
			c
			e
Serie - titulo	<u>440</u>	---	a
			n
			p
			v
Notas gerais	<u>500</u>	---	a
Nota biblioq.	<u>504</u>	---	a
Nota conteúdo	<u>505</u>	<u>0</u>	a
Não existe	<u>590</u>	---	a
Assunto-nome	<u>600</u>	<u>?</u>	a
			b
			c
			d
			e
			f
			x
			y
			z
Assunto-corp.	<u>610</u>	<u>?</u>	a
			b
			c
			d
			e
			f
			x
			y
			z
Assunto-evento	<u>611</u>	<u>?</u>	a
			c
			d
			n
			x
			y
			z

Assunto t3pico	<u>650</u>	<u>4</u>	a	_____
			x	_____
			y	_____
			z	_____
Assunto-geogr.	<u>651</u>	<u>4</u>	a	_____
			x	_____
			y	_____
			z	_____
C3d. As. CAPES	<u>697</u>		a	_____
Ent.sec.-nome	<u>700</u>	<u>?</u>	a	_____
			b	_____
			c	_____
			d	_____
			e	_____
Ent.sec.-corp.	<u>710</u>	<u>1</u>	a	_____
			b	_____
			c	_____
			d	_____
Ent.sec.-event.	<u>711</u>	<u>2</u>	a	_____
			b	_____
			c	_____
Ent.sec.-tit.a.	<u>740</u>	<u>?</u>	a	_____
			b	_____
			c	_____
			d	_____
			e	_____

FIGURA 4: Formato de entrada

FONTE: Banco de Dados ATHENA da UNESP, 2003.

Vale salientar que as figuras que aqui ir3o ilustrar alguns exemplos, no decorrer deste cap3tulo, foram retiradas do Banco de Dados ATHENA da UNESP que utiliza o software ALEPH 500, da Base de dados BIBLIODATA que utiliza o software VTLS e da Base de dados da Library of Congress – LC, portanto, essas imagens do formato MARC21 vistas de outros softwares que n3o sejam estes poder3o sofrer algumas altera33es.

O Formato de Armazenamento Interno diz respeito a aloca33o dos dados na mem3ria do computador. Suas caracter3sticas mais desej3veis s3o efici3ncia e economia no armazenamento e recupera33o dos dados (BALBY, 1995).

O Formato de Interc3mbio ou comunica33o 3 gerado por programa, a partir do formato de armazenamento interno, com o fim de possibilitar a incorpora33o e interpreta33o de dados bibliogr3ficos por sistemas e computadores diferentes. 3 voltado para a comunica33o de dados bibliogr3ficos entre computadores, e suas caracter3sticas mais desej3veis s3o a efici3ncia e economia na troca de dados. Um formato de interc3mbio deve permitir a comunica33o de dados entre computadores diferentes com o m3nimo disp3ndio em reformata33o dos dados (BALBY, 1995).

Desta forma, um registro no formato de intercâmbio ou comunicação é manipulado somente pelo programador e o computador, mas não deixa de ser interessante ou porque não dizer importante entender como o diretório funciona.

Observe as figuras 5 e 6 abaixo :

```
00644nam 2200217 a
4500001001200000008004100012020002400053040001800077080001400095090002000109100001900129245009200
1482600031002403000012002714900032002835040018003156500018003336500031003516500028003829980016004
10ER000295397990304s1997 po 000 0 por d a9720341246 (broch.)
aBIRjFGVBbpor a301.173.7 a301.173.7bD813-1 aDubar, Claude.12aA socialização
:bconstrução das identidades sociais e profissionais /cClaude Dubar. -- aPorto :bPorto
Ed.,c1997. a240p. --0 a(Ciencias da educação;v24) aBibliografia. 4aSocializacao.
4aSocializacao profissional. 4aIdentidade (Psicologia) aDR ER UP UV
```

FIGURA 5: Formato de intercâmbio

FONTE: Base de dados BIBLIODATA em CD-ROM, 2003.

```
00644nam 2200217 a
4500001001200000008004100012020002400053040001800077080001400095090002000109100001900
1292450092001482600031002403000012002714900032002835040018003156500018003336500031003
51650002800382998001600410-ER000295397-990304s1997 po 000 0 por d- a9720341246
(broch.)- aBIRjFGVBbpor- a301.173.7- a301.173.7bD813-1 aDubar, Claude.-12aA socialização :-
bconstrução das identidades sociais e profissionais /cClaude Dubar. -- aPorto :bPorto Ed.,c1997.- a240p.
--0 a(Ciencias da educação;v24)- aBibliografia.- 4aSocializacao.- 4aSocializacao profissional.- 4-
aIdentidade (Psicologia)- aDR ER UP UV-
```

FIGURA 6: cópia da figura 5 editada com cores para melhor visualizar o diretório.

As primeiras 24 posições constituem o líder. Neste exemplo, o líder ocupa a primeira linha e o início da segunda e termina com “4500”. Imediatamente após o líder, começa o Diretório. A primeira etiqueta é a 001. Após cada etiqueta, as quatro posições seguintes mostram o tamanho do campo. Os dados no campo 001 (número de controle) neste registro tem 12 caracteres de extensão. As cinco posições seguintes mostram o ponto de início deste campo dentro da seqüência de caracteres que segue o diretório. O campo 001 começa na posição 00000 (a primeira posição é a posição 0). A etiqueta seguinte é a 008, que tem quarenta e um caracteres de extensão e começa na décima segunda posição (o tamanho da posição anterior 12 somado à posição de início – 00000 – é igual a 12). A etiqueta seguinte é

a 020. Possui 24 caracteres de extensão e começa na quinquagésima terceira posição (41 + 12 = 53) (FURRIE, 2000).

Os finalizadores de campo que podem ser vistos na figura 5 capturada da Base BIBLIODATA como um (■) marcam o fim do diretório e o fim de cada campo que se segue. O finalizador de registro (mostrado como (■ ■) neste exemplo) termina cada registro bibliográfico. Não esquecendo que os espaços são contados, bem como os finalizadores de campo citados. E duas posições de caracteres são sempre reservadas para indicadores no início de cada campo. (FURRIE, 2000).

As cores dos caracteres dentro da figura 6 significam:

Preto(Negrito) = Líder ;

Azul = Etiquetas MARC21 ;

Vermelho = As quatro posições mostram o tamanho da etiqueta MARC21;

Verde = As cinco posições indica o ponto de início da etiqueta dentro da seqüência de caracteres que segue o diretório.

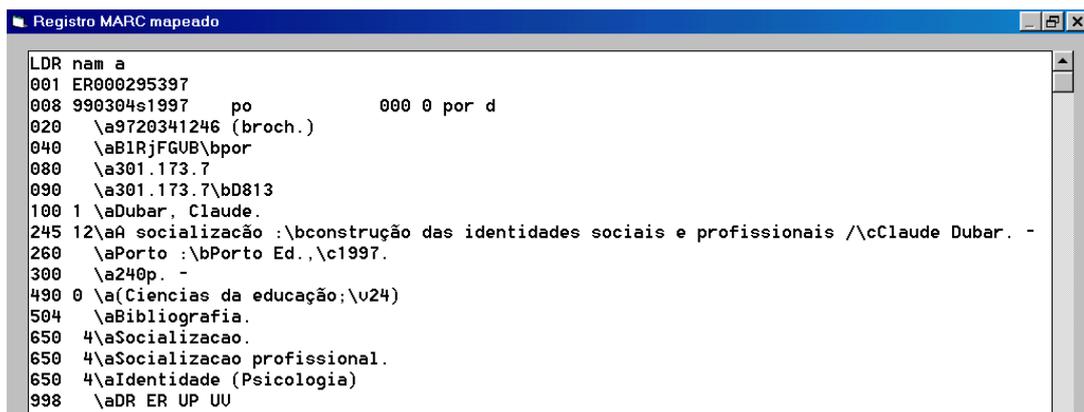
Este diretório contém as suas etiquetas MARC21 organizadas da seguinte forma:

Etiqueta	Tamanho	Início
001	0012	00000
008	0041	00012
020	0024	00053
040	0018	00077
080	0014	00095
090	0020	00109
100	0019	00129
245	0092	00148
260	0031	00240
300	0012	00271
490	0032	00283
504	0018	00315
650	0018	00333
650	0031	00351
650	0028	00382
998	0016	00410

FIGURA 7: Etiquetas MARC21 organizadas em forma de tabela

“Observe que a soma da segunda e terceira colunas em qualquer linha é igual ao número da terceira coluna na linha seguinte. O ponto de início de um campo mais a sua extensão é igual ao ponto de início do campo seguinte” (FURRIE, 2000, p. 73). Lembrando que, tanto o tamanho como o início de cada etiqueta variam de registro para registro devido as especificidades de cada item documentário, assim como algumas etiquetas também. Por exemplo: Uma determinada obra que tem um organizador, não tendo portanto um responsável intelectual para utilizar a etiqueta 100 (AUTOR), terá a entrada principal da obra no registro MARC21 pela etiqueta 245 (TÍTULO), acrescido de uma etiqueta 700 (ENTRADA SECUNDÁRIA) de nome pessoal. Portanto, este registro iria se diferenciar do exemplo citado, com algumas etiquetas diferentes, assim como o tamanho e início de cada uma delas.

A figura 8 mostra o registro citado em formato de intercâmbio, agora na tela de exibição com etiquetas MARC21. Esta forma de exibição é proporcionada pelo software do sistema, e é vista e utilizada geralmente pelo profissional bibliotecário.



```

LDR nam a
001 ER000295397
008 990304s1997 po 000 0 por d
020 \a9720341246 (broch.)
040 \aB1RjFGUB\bp
080 \a301.173.7
090 \a301.173.7\bd813
100 1 \aDubar, Claude.
245 12\aa socialização :\bconstrução das identidades sociais e profissionais /\cClaude Dubar. -
260 \aPorto : \bPorto Ed., \c1997.
300 \a240p. -
490 0 \a(Ciências da educação;\v24)
504 \aBibliografia.
650 4\asocializacao.
650 4\asocializacao profissional.
650 4\aidentidade (Psicologia)
998 \aDR ER UP UU
  
```

FIGURA 8 : Exibição com etiqueta

FONTE: Base de dados BIBLIODATA em CD-ROM, 2003.

As telas que são exibidas para os usuários dos catálogos online de acesso público, ou seja, os OPACs são telas de exibição formatadas, visto que as etiquetas MARC21 são incompreensíveis para o público em geral. Nos programas para catálogos online existe um

subprograma que formata cada registro para melhor apresentação dos dados para o usuário do catálogo online (FURRIE, 2000). Veja a figura 9 a seguir:



FIGURA 9 : Exibição formatada

FONTE: Base de dados BIBLIODATA Online, 2003.

3.5 Comunicação entre os Designadores de Conteúdos

As etiquetas MARC21 e os seus respectivos indicadores e subcampos, comunicam-se entre si devido a estrutura do formato e desta forma, ajudam o bibliotecário catalogador a fazer a entrada correta de um registro em seu Banco de dados. Observe a figura 10:

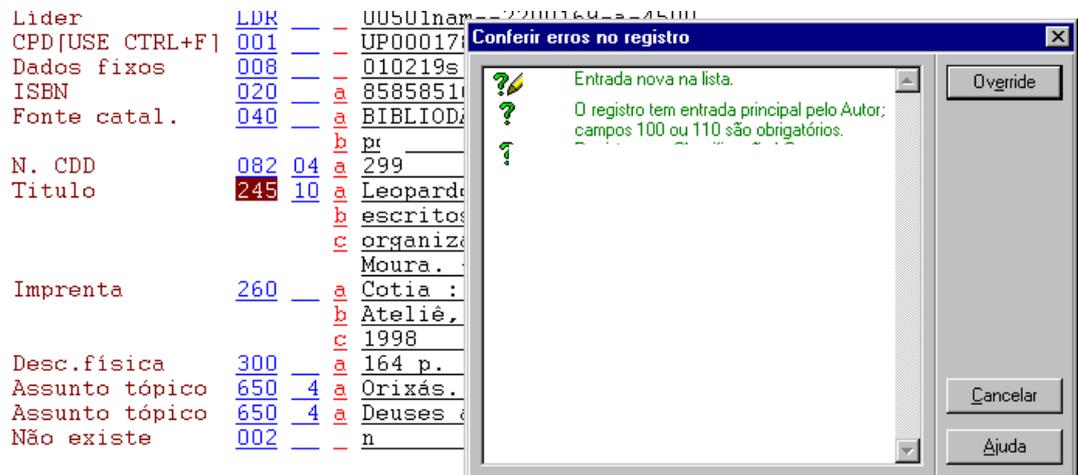


FIGURA 10: Campo 245 - indicador 1

No exemplo da figura 10 podemos perceber que este registro está sendo adicionado em um banco de dados pelo catalogador, mas ao confirmar a entrada de dados o sistema retornou com uma mensagem dizendo: O registro tem entrada principal pelo Autor : campos 100 ou 110 são obrigatórios. Mas como o sistema sabe disso? Simples. No campo 245 o catalogador não percebeu que o indicador 1 estava igual a 1, que quer dizer que o registro precisa ter um responsável intelectual, e portanto, campos 100 ou 110.

Na figura 11, o catalogador colocou o indicador 1 do campo 245 igual a zero e veja novamente a mensagem do sistema:

Líder	<u>LDR</u>	—	—	00501nam--2200169-a-4500
CPD[USE CTRL+F]	<u>001</u>	—	—	UP00017
Dados fixos	<u>008</u>	—	—	010219s
ISBN	<u>020</u>	—	<u>a</u>	8585851
Fonte catal.	<u>040</u>	—	<u>a</u>	BIBLIOD
			<u>b</u>	por
N. CDD	<u>082</u>	<u>04</u>	<u>a</u>	299
Titulo	<u>245</u>	<u>00</u>	<u>a</u>	Leopard
			<u>b</u>	escrito
			<u>c</u>	organiz
				Moura.
Imprensa	<u>260</u>	—	<u>a</u>	Cotia :
			<u>b</u>	Ateliê,
			<u>c</u>	1998
Desc.física	<u>300</u>	—	<u>a</u>	164 p.
Assunto tópico	<u>650</u>	<u>4</u>	<u>a</u>	Orixás.
Assunto tópico	<u>650</u>	<u>4</u>	<u>a</u>	Deuses
Não existe	<u>002</u>	—	—	n

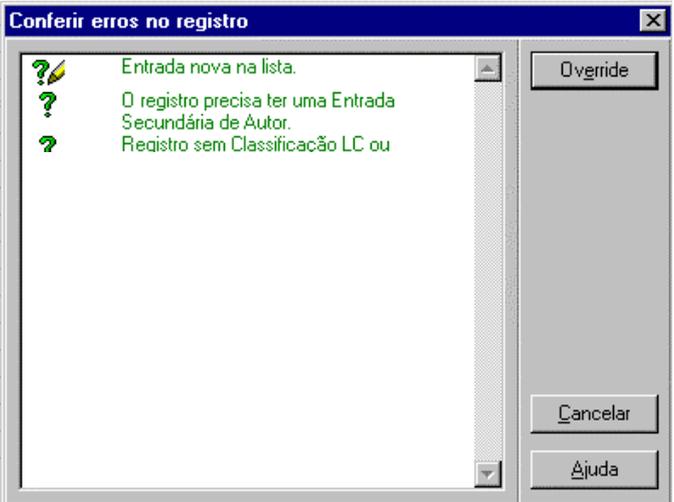


FIGURA 11: Campo 245 – indicador 1 igual a 0.

Neste exemplo o catalogador corrigiu o campo 245 e seus indicadores 1 e 2, que respectivamente significam: Não tem responsável intelectual e não tem caracteres a desprezar. Só que o indicador 1 igual a zero, está comunicando ao sistema que este registro precisa de uma entrada secundária, o campo 700.

Na figura 12 podemos observar que o catalogador acrescentou o campo 700 e inseriu o registro no banco de dados sem problemas.



FIGURA 12: Registro com entradas corretas.

A mesma comunicação ocorre com o campo 008 e o campo 260, mais especificamente com o subcampo “c” deste campo.

Quando o catalogador aproveita um registro catalográfico e bibliográfico seja da sua própria instituição ou de outra, o primeiro passo a se fazer é mudar o número de controle do registro (001), ou seja o RG do registro, pois este número é único, individual para cada registro, assim como nossas carteiras de identidade.

Por exemplo, um catalogador fictício recebeu em sua instituição a obra do Autor Antonio Augusto Faria; Título: Getúlio Vargas e sua época; Edição: 9ª ; Local: São Paulo; Editora: Global; Ano: 2001. Então ele faz uma consulta em seu banco de dados e percebe que ele tem catalogado uma edição anterior desta obra, mudando somente a edição e o ano. Como as tecnologias estão aí para facilitar a nossa vida e não complicá-la, este catalogador resolve então aproveitar esse registro para criar um novo, mas somente depois de confirmar que em

outras bases além da sua não há a obra catalogada, para isso será preciso mudar três coisas, o número de controle do registro, a edição e o ano.

Veja a figura 13: representa o registro que o catalogador tem em sua base de dados e o qual ele irá aproveitá-lo para criar um novo registro.

```

Lider          LDR  -- 00730nam--2200229-a-4500
CPD[USE CTRL+F] 001 -- PR000207312
Dados fixos     008 -- 991104s1997----spbac-----000-0dpor-d
ISBN           020 -- 8526000713 (broch.)
Fonte catal.   040 -- BIRjFGVB
                b pcr
N. CDD         082 04 -- 981.05
Não existe     090 -- 981.05
                b F224
                c 8.ed.
Nome pessoal   100 1  a Faria, Antonio Augusto,
                d 1949-
Titulo        245 10 a Betulio Vargas e sua epoca /
                c Antonio Augusto Faria, Edgard Luiz de Barros.-
Edição        250 -- 8.ed.-
Imprenta      260 -- São Paulo :
                b Global,
                c 1997.
Desc.fisica   300 -- 108p. :
                b il., retrs.
Serie - titulo 440 1  a (Historia popular;
                v n.8)
Nota bibliog. 504 -- Inclui bibliografia
Assunto-geogr. 651 4  a Brasil
                x Historia -Crise de 1954.
Assunto-geogr. 651 4  a Brasil
  
```

FIGURA 13: Registro aproveitável

Observe agora a figura 14 : O catalogador já fez uma cópia do registro representado na figura 13, acessou o campo 001 e está fazendo a primeira modificação: alterando o número de controle para assim poder criar um novo registro.

Formulário para campo 001 - Formato BK

Número de CPD (NR) PR000207312.

OK

Ajuda no campo

Cancelar

Ajuda

```

Lider          LDR  -- 00730nam--2200229-a-4500
CPD[USE CTRL+F] 001 -- PR000207312
                x 8526000713 (broch.)
                x 8.ed.-
                x São Paulo :
                x Global,
                x 1997.
                x 108p. :
                x il., retrs.
                x (Historia popular;
                x n.8)
                x Inclui bibliografia
                x Brasil
                x Historia -Crise de 1954.
                x Brasil
  
```

FIGURA 14: Registro sendo modificado.

Na figura 15 o catalogador está inserindo um novo número de controle ao registro, gerado pela instituição.

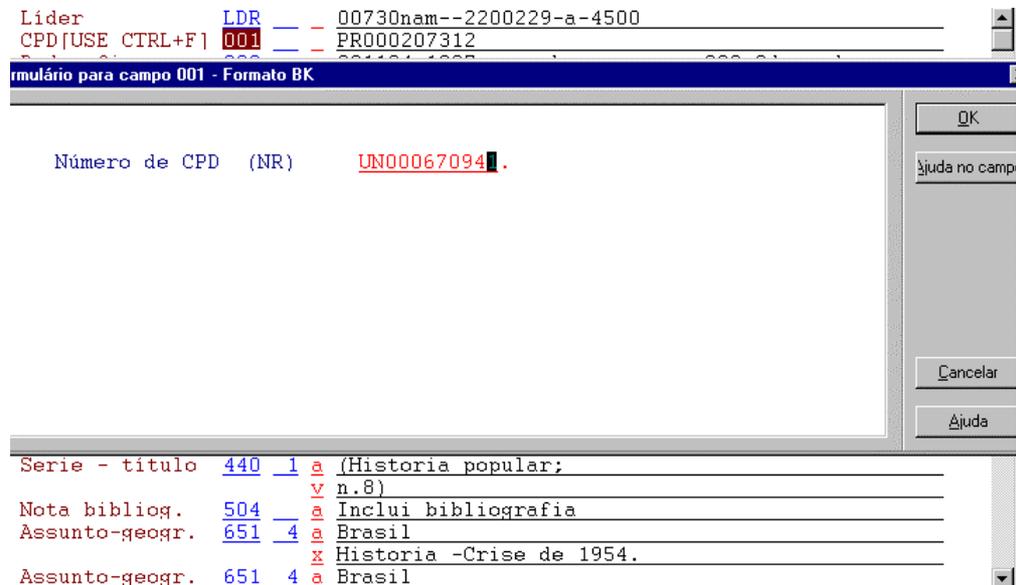


FIGURA 15 : Novo número de controle adicionado a cópia do registro

Na figura 16 o catalogador alterou o campo 001, o subcampo "c" do campo 260 e corrigiu o campo 245. Observe que o campo 008 continua da mesma forma como o registro original.

The screenshot shows a MARC record with the following fields:

Líder	LDR			00730nam--2200229-a-4500
CPD[USE CTRL+F]	001			UN000670941
Dados fixos	008			991104s1997---spbac-----000-0dpor-d
ISBN	020		a	8526000713
Fonte catal.	040		a	BIBLIODATA
			b	por
			a	x5x5
Per. cron.[F2]	045		a	981.05
N. CDD	082	04	a	
Nome pessoal	100	1	a	Faria, Antonio Augusto,
			d	1949-
Título	245	10	a	Getúlio Vargas e sua época /
			c	Antonio Augusto Faria, Edgard Luiz de Barros. -
Edição	250		a	9. ed. -
Imprenta	260		a	São Paulo :
			b	Global,
			c	2001
Desc.fisica	300		a	108 p. :
			b	il., retrs. -
Serie - título	440	1	a	(História popular ;
			v	n. 8)
Nota biblioq.	504		a	Inclui bibliografia
Assunto-geogr.	651	4	a	Brasil -
			x	Historia -
			y	Crise de 1954.
Assunto-geogr.	651	4	a	Brasil -

FIGURA 16: Registro alterado

A figura 17 mostra que o sistema está pedindo uma confirmação do catalogador para salvar o registro no banco de dados. Veja que o campo 008 continua com a mesma data do registro original 1997.

Lider	LDR	--	00730nam--2200229-a-4500
CPD[USE CTRL+F]	001	--	UN000670941
Dados fixos	008	--	991104s1997--spbac-----000-0dpor-d
ISBN	020	--	a 8526000713
Fonte catal.	040	--	a BIBLIODATA
		b	por
Per. cron.[F2]	045	--	a x5x5
N. CDD	082	04	a 981.05
Nome pessoal	100	1	a Faria, Antor
		d	1949-
Titulo	245	10	a Getúlio Vars
		c	Antonio Augu
Edição	250	--	a 9. ed. -
Imprenta	260	--	a São Paulo :
		b	Global,
		c	2001
Desc.fisica	300	--	a 108 p. :
		b	il., retrs.
Serie - titulo	440	1	a (História popular ;
		v	n. 8)
Nota bibliog.	504	--	a Inclui bibliografia
Assunto-geogr.	651	4	a Brasil -
		x	Historia -
		y	Crise de 1954.
Assunto-geogr.	651	4	a Brasil -

Salvar servidor/drive local

Salvar registro na biblioteca UEP01 como um novo registro

Continuar?

Trocar nível do catalogador

Continuar

Cancelar

Ajuda

FIGURA 17: Registro sendo confirmado

O catalogador ao confirmar a entrada do registro no banco de dados bibliográficos, automaticamente o sistema alterou o campo 008 como podemos observar na figura 18, poupando o bibliotecário de acessar o referido campo somente para mudar a data.

Lider	LDR	--	00730nam--2200229-a-4500
CPD[USE CTRL+F]	001	--	UN000670941
Dados fixos	008	--	991104s2001--spbac-----000-0dpor-d
ISBN	020	--	a 8526000713
Fonte catal.	040	--	a BIBLIODATA
		b	por
Per. cron.[F2]	045	--	a x5x5
N. CDD	082	04	a 981.05
Nome pessoal	100	1	a Faria, Antonio Augusto,
		d	1949-
Titulo	245	10	a Getúlio Vargas e sua época /
		c	Antonio Augusto Faria, Edgard Luiz de Barros. -
Edição	250	--	a 9. ed. -
Imprenta	260	--	a São Paulo :
		b	Global,
		c	2001
Desc.fisica	300	--	a 108 p. :
		b	il., retrs. -
Serie - titulo	440	1	a (História popular ;
		v	n. 8)
Nota bibliog.	504	--	a Inclui bibliografia
Assunto-geogr.	651	4	a Brasil -
		x	Historia -
		y	Crise de 1954.
Assunto-geogr.	651	4	a Brasil -
		x	Historia -

FIGURA 18: Registro salvo corretamente no banco de dados

3.6 Relações existentes entre etiquetas

Assim como há uma comunicação entre os designadores de conteúdo, entre algumas etiquetas existe uma relação entre elas. Neste caso o sistema não acusa a falta da etiqueta no registro, e ele é salvo com ou sem a etiqueta relacionada e desta forma o registro fica com informações incompletas.

Algumas etiquetas relacionadas são:

043 Código de área geográfica

045 Código de período cronológico

650 Assunto – Termo tópico

651 Assunto – Nome geográfico

Veja os exemplos a seguir:

Lider	LDR	—	—	00619nam--2200193-a-4500
CPD[USE CTRL+F]	001	—	—	BN000127302
Dados fixos	008	—	—	850315s1984----spb-----r----s000-0-por-d
Fonte catal.	040	—	a	BIBLIODATA
			b	por
Area geog. [F8]	043	—	a	s-bl---
Per. cron.[F2]	045	—	a	w0w9
N. CDD	082	04	a	709.8161
Titulo	245	00	a	Artes no Brasil no século XIX :
			b	um ciclo de palestras, agosto-setembro de 1977. -
Imprensa	260	—	a	São Paulo :
			b	Pinacoteca do Estado,
			c	[1984?]
Desc.fisica	300	—	a	52 p.
Notas gerais	500	—	a	Dados retirados da capa
Notas gerais	500	—	a	Inclui bibliografia
Assunto tópico	650	4	a	Artes -
			z	Brasil -
			x	História -
			y	Século XIX - Discursos, ensaios, conferências.
Ent.sec.-corp.	710	2	a	Pinacoteca do Estado (SP)

FIGURA 19 : Etiquetas relacionadas

No exemplo acima (FIG. 19) a etiqueta 650 está diretamente relacionada com as etiquetas 043 e 045 da seguinte forma:

O subcampo “z” do campo 650 indica a área geográfica do assunto Artes como Brasil e a etiqueta relacionada é a etiqueta 043 que indica a área geográfica Brasil utilizando

um código de 7 caracteres. Existe uma tabela chamada Códigos de Área Geográfica - GACs que lista os códigos de todos os países, estados e cidades, ou seja, de todas as áreas geográficas. Veja a figura 20:

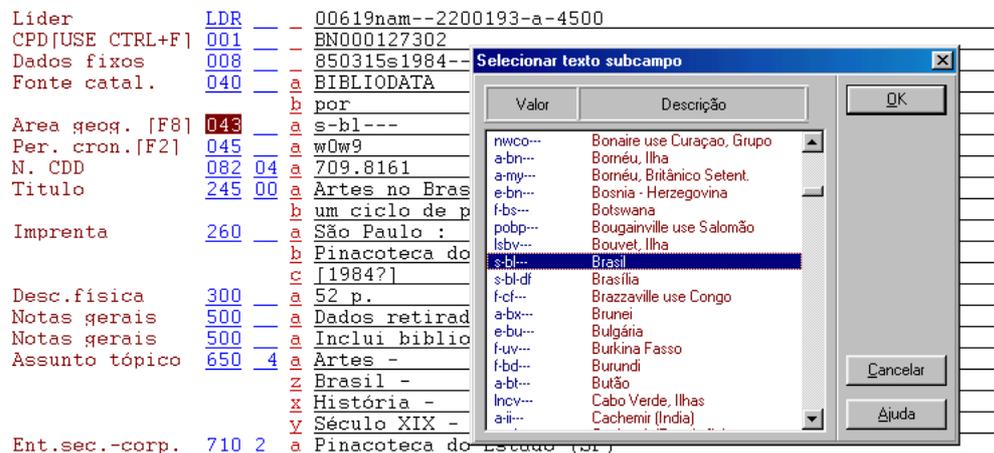


FIGURA 20: Tabela da etiqueta 043

O subcampo “y” do campo 650 indica a época do assunto Artes como século XIX, a etiqueta relacionada é a etiqueta 045 que indica o século em código de quatro dígitos. Existe a Tabela de Código de Período Cronológico que lista os códigos de todas as épocas (período de tempo) antes e depois de Cristo. Veja a figura 21:



FIGURA 21: Tabela da etiqueta 045

No exemplo da (FIG.21) a Tabela de Código de Período Cronológico codifica uma letra do alfabeto para representar cada século. Para o século XIX por exemplo é utilizado a letra W. A codificação da etiqueta 045 ficou w0w9 que significa:

W – representa o século

0 – representa a década do século

W- representa o século

9 – representa o ano do século

3.7 **Campo 856**

A inclusão do campo 856 na estrutura do **MARC21** foi de iniciativa da Library of Congress - LC em 1992 com a intenção de adaptá-lo a utilização de recursos disponíveis na Internet.

Este campo permite criar uma ligação entre um registro bibliográfico e um recurso remoto disponível em formato eletrônico. A informação contida neste campo é suficiente para permitir a transferência eletrônica do arquivo, assinatura de um jornal eletrônico, ou ligação a um catálogo de biblioteca.

O primeiro indicador define o método de acesso ao recurso eletrônico como, e-mail, FTP, Dial-up, HTTP e outros. Se este recurso é acessível por mais de um método, este se repete.

O segundo indicador indica a relação entre o recurso eletrônico identificado no campo 856 e o documento descrito no registro como um todo, por exemplo, tipo de fonte, versão da fonte, fonte relacionada e outros.

Neste exemplo da (FIG.22) temos um registro que utiliza o campo 856.

What minds can do : intentionality in a non-intentional world / Pierre Jacob.

LC Control Number: **96014157**

000 01274cam 2200313 a 450
 001 3051765
 005 20020829202144.0
 008 960321s1997 enka b 001 0 eng
 035 __ |9 (DLC) **96014157**
 906 __ |a 7 |b cbc |c orignew |d 1 |e ocip |f 19 |g y-gencatlg
 955 __ |a pc14 to sa00 03-21-96; sh20 03-22-96; sh08 03-29-96; sh15 03-29-96; CIP ver. pv07 04-14-97
 010 __ |a **96014157**
 020 __ |a 0521574013 (hardback)
 020 __ |z 0521474366 (pbk.)
 040 __ |a DLC |c DLC |d DLC
 050 00 |a B105.I56 |b J33 1997
 082 00 |a 128/.2 |2 20
 100 1_ |a Jacob, Pierre, |d 1949-
 245 10 |a What minds can do : |b intentionality in a non-intentional world / |c Pierre Jacob.
 260 __ |a Cambridge ; |a New York : |b Cambridge University Press, |c 1997.
 300 __ |a xii, 299 p. : |b ill. ; |c 23 cm.
 440 _0 |a Cambridge studies in philosophy
 504 __ |a Includes bibliographical references (p. 281-294) and index.
 650 _0 |a Intentionality (Philosophy)
 650 _0 |a Philosophy of mind.
 856 42 |3 Publisher description |u <http://www.loc.gov/catdir/description/cam027/96014157.html>
 856 41 |3 Table of contents |u <http://www.loc.gov/catdir/toc/cam024/96014157.html>
 920 __ |a ****LC HAS REQ'D # OF SHELF COPIES****
 991 __ |b c-GenColl |h B105.I56 |i J33 1997 |t Copy 1 |w BOOKS

FIGURA 22: Registro utilizando o campo 856
FONTE: Catálogo online da Library of Congress - LC.

3.8 Normas adotadas pelo MARC21

Foi considerado no projeto do formato a adoção de normas internacionais de descrição bibliográfica e de ferramentas de normalização.

O sistema adota para a descrição bibliográfica o Anglo-American Cataloguing Rules–2ª edição - AACR2 e para normalização utiliza a norma internacional de padronização

ISO 2709 e a norma **ANSI Z39.2**, e para recuperação e intercâmbio de dados catalográficos e bibliográficos em ambiente automatizado, o Protocolo **Z39.50** descritos a seguir.

3.8.1 Anglo- American Cataloguing Rules – 2ª edição - AACR2

O **Anglo-American Cataloguing Rules – 2ª edição - AACR2**, ou seja, Código de Catalogação Anglo-Americano – 2ª edição, é uma obra fundamental e imprescindível nas unidades de informação, sejam elas bibliotecas, centros de documentação e outros. O código **AACR2** é resultado de trabalho cooperativo de várias entidades representativas no panorama biblioteconômico mundial.

Algumas Associações mais representativas são: **American Library Association - ALA**, **Library of Congress - LC**, **British Library Association**, **Canadian Library Association** e outras (LEHNUS, 1971).

O **AACR2** foi publicado em 1978 e revisado em 1988 em edições Britânicas e Americanas.

Foi publicado no Brasil em 1983 em dois volumes. O volume 1 relativo a descrição bibliográfica e o volume 2 para cabeçalhos, títulos uniformes e remissivas.

Para El-Sherbini (2000, p.181, tradução nossa) o **AACR2** “é uma ferramenta projetada para ajudar a organizar o acesso aos dados. É um código de regras utilizado para a criação de descrições bibliográficas de documentos, um sistema de classificação criado simplesmente para prover acesso aos materiais”.

Atualmente está sendo publicada a nova edição do **AACR2**, e certamente será a nova edição adotada para a descrição bibliográfica.

3.8.2 Norma ISO 2709

“Uma norma é uma fórmula que tem valor de regra, em geral indicativa e algumas vezes imperativa. Ela define as características que deve ter um objeto e as suas características de uso, bem como as características de um procedimento e /ou de um método”(GUINCHAT e MENO, 1994, p. 433). Uma norma é adotada como base ou medida para a realização ou avaliação de algo (FERREIRA, 1993).

A normalização para os dados catalográficos e bibliográficos é fundamental para que possa haver o compartilhamento e a cooperação entre várias unidades de informação, algo indispensável hoje em dia, possibilita a simplificação e racionalização dos métodos e técnicas utilizadas padronizando os dados informacionais e desta forma diminuindo custos e tempo para efetuar o intercâmbio de informações.

A **International Standardization Organization - ISO**, ou seja, Organização Internacional para Padronização é o organismo principal de normalização mundial e sua atividade se estende a todos os campos. É uma organização não governamental estabelecida em 1947.

O objetivo da **ISO** é promover o desenvolvimento da padronização e atividades relacionadas no mundo inteiro, para facilitar a troca internacional de bens e serviços, e a cooperação nas esferas intelectual, científica, tecnológica e econômica (**ISO**, 2002). Assim como a World Wide Web Consortium – W3C, organização criada em 1994 e que tem por objetivo padronizar e normatizar os padrões para a Web, conduzindo-a a seu potencial máximo e garantindo a sua interoperabilidade (W3C, 1994-2002).

O trabalho da **ISO** resulta em acordos internacionais que são publicados como Padrões Internacionais. Padrões são documentos gerados a partir de acordos e que contêm especificações técnicas ou outros critérios que são precisos para serem usados constantemente

como regras, diretrizes, ou definições de características, para assegurarem que materiais, produtos, processos e serviços são ajustados, padronizados para servir a um determinado propósito (ISO, 2002).

Por exemplo, o formato dos cartões de crédito e cartões telefônicos que se tornaram comuns, é derivado de uma norma internacional padrão ISO. Aderindo ao padrão que define tal característica como uma espessura de (0,76 mm), os cartões podem ser usados mundialmente (ISO, 2002).

Portanto, os padrões internacionais contribuem para tornar a vida mais simples, e aumentar a confiança e efetividade dos bens e serviços que todos nós usamos (ISO, 2002).

Com este mesmo propósito se aplica a norma ISO 2709 para a normalização do formato MARC21. A norma internacional de padronização ISO 2709 (Documentation Format for Bibliographic Interchange on Magnetic tape), foi desenvolvida pelo Comitê Técnico ISO/TC46, encarregado da elaboração de normas no campo da documentação, juntamente com subcomitê SC4- encarregado da automação em documentação.

A norma ISO 2709 especifica os requisitos para o formato de intercâmbio de registros bibliográficos que descrevam todas as formas de documentos sujeitos a descrição bibliográfica. Não define a extensão do conteúdo de documentos individuais, nem designa significado algum para os parágrafos, indicadores ou identificadores, sendo essas especificações as funções dos formatos de implementação.

A forma como os dados estão estruturados em meio magnético possibilitam o intercâmbio de registros bibliográficos. Mesmo assim, esta característica não elimina a incompatibilidade entre os registros que utilizam diferentes formatos de entrada e, principalmente, diferentes regras de entrada de dados.

A estrutura do formato, mesmo quando semelhante, não garante a compatibilidade do registro. É preciso considerar o conteúdo dos registros que irão influenciar fortemente na recuperação da informação.

A **ISO 2709** apresenta uma estrutura geral, ou seja um molde ou esqueleto projetado especialmente para a comunicação entre sistemas de processamento de dados e não para uso como formato de processamento dentro dos sistemas (CÔRTE et al., 2000).

3.8.3 Norma ANSI Z39.2

O American National Standards Institute - **ANSI**, é uma organização privada sem fins lucrativos que administra e coordena a padronização norte-americana, fundada em 1918.

O objetivo do Instituto é aumentar a competitividade global de negócios dos Estados Unidos e a qualidade de vida norte-americana promovendo padrões e seus sistemas de avaliação (NISO, 1994).

A norma **ANSI Z39.2** especifica as exigências para o formato de intercâmbio de informação generalizada que acomodam muitos tipos de dados, especialmente descrição bibliográfica de todos os tipos de materiais e dados relacionados como autoridade, cabeçalhos de assunto, circulação e outros.

Descreve uma estrutura geral, ou seja, uma forma especificamente projetada para troca de dados entre sistemas de comunicação de dados e não necessariamente para uso como um formato de processamento dentro dos sistemas.

Não especifica o conteúdo de um registro e não atribui significado as etiquetas, indicadores, ou identificadores de elementos dos dados. Essas especificações serão providas por formatos particulares de implementação.

O formato pode ser usado para o intercâmbio de registros usando vários meios de comunicação (NISO, 1994).

3.8.4 Protocolo de comunicação Z39.50

No contexto da Internet, protocolo é um conjunto de regras que os computadores utilizam para se comunicarem entre si através das redes. A Internet por exemplo, não seria hoje uma realidade sem o protocolo **TCP-IP** (Transmission Control Protocol/Internet Protocol) que possibilita a comunicação entre os computadores. Da mesma forma, sem o protocolo **Z39.50** não seria possível acessar simultaneamente os catálogos online de várias bibliotecas via Internet.

Protocolo é conceituado por alguns autores da seguinte forma:

Para Brookshear (2000, p. 133),

Os conjuntos de regras que administram a comunicação entre os diferentes componentes de um sistema computacional são denominados “PROTOSCOLOS”, em alusão aos protocolos usados na sociedade para administrar as relações humanas. Por meio dos protocolos da rede é que se definem os detalhes de cada atividade, incluindo o modo como são enviadas as mensagens, a maneira como a autorização para transmitir mensagens é delegada às máquinas, e a forma como são manipuladas as tarefas de empacotar e desempacotar mensagens para a transmissão. (grifo do autor).

Para Krol (1993) citado por Ferreira (1994, p. 258),

Protocolo é um conjunto de regras e comandos preestabelecidos entre os componentes de uma rede para que se efetue suas comunicações, ou seja, para que os computadores envolvidos “falem” entre si. Os protocolos são, geralmente, documentos em páginas e páginas de material impresso, porém,

como ocorre com tantas outras tecnologias complexas, o usuário não precisa conhecê-los em detalhes para que possa deles se utilizar.

“Protocolo de comunicação é um elenco de regras ou padrões cuja finalidade é permitir que os computadores se interliguem e troquem informações com o menor número de erros possível” (ROSSETTO, 1997, p.1).

O protocolo **Z39.50** foi desenvolvido num projeto conhecido como **Linked Systems Project – LSP**, no início de 1980 . Este projeto envolveu a **Library of Congress - LC** (Biblioteca do Congresso Norte Americano), **Online Computer Library Center - OCLC**, **RLIN** e a **Western Library Network - WLN**. O projeto **LSP** construiu uma rede experimental para transferir registros bibliográficos entre os participantes e desenvolveu um precursor do protocolo **Z39.50**.

Em 1984 o protocolo desenvolvido pelo **LSP** foi enviado para a **National Information Standards Organization - NISO**, **American National Standards Institute – ANSI**, publicadores e provedores de serviços de informação para adicionais desenvolvimento e implementações e assim ser considerado um padrão nacional dos Estados Unidos.

A primeira versão do protocolo **Z39.50** foi adotada como um padrão nacional dos Estados Unidos em 1988.

Depois da primeira versão surgiram outras com várias modificações e implementações. Surgiu então a necessidade de se adotar somente uma versão do protocolo. As instituições envolvidas entraram em um consenso e em 1998 foi aprovada oficialmente a versão **Z39.50** de 1995 normalizada pela **ISO 23950**, como padrão internacional.

O protocolo **Z39.50** embora originalmente tenha surgindo no contexto das bibliotecas e do mundo bibliográfico, não há nada no protocolo **Z39.50** que restrinja seu uso somente a esses tipos de dados. Ele foi especificamente projetado para ser um mecanismo capaz de manipular qualquer tipo de dados como, imagens, documentos de texto completo,

dados bibliográficos e catalográficos e outros, permitindo a pesquisa e recuperação dessas informações em redes de computadores distribuídos (NEEDLEMAN, 2001).

O protocolo **Z39.50** utiliza-se de um modelo cliente/servidor para possibilitar a busca e recuperação de informações na Internet. Permite a um usuário de um sistema pesquisar e recuperar informações de outros sistemas. Para Abiteboul, Buneman e Suciú (2000) uma arquitetura cliente/servidor funciona da seguinte forma: O cliente (uma pessoa ou um programa) emite uma consulta que é processada, compilada em um código otimizado e executada, e dados de resposta são devolvidos pelo servidor.

Para Côrte et al (1999, p. 247) “o padrão Z39.50 é um padrão de “middleware” (sic) cliente-servidor”. Segundo Abiteboul, Buneman e Suciú (2000, p. 5) “middleware é o software que transforma, integra ou senão adiciona valor aos dados”. Ficando entre a comunicação do cliente com o servidor.

Com o uso deste padrão é possível a interoperabilidade, ou seja, a compatibilidade entre vários sistemas diferentes. permite-se que determinado servidor possa prestar serviços para diversos clientes ou fornecedores, e um determinado servidor possa usar os serviços de diversos servidores ou fornecedores diferentes. Côrte et al (1999, p. 247) esclarece melhor com o seguinte exemplo:

[...] se este protocolo não existisse, a biblioteca que quisesse recuperar registros bibliográficos de vários bancos de dados deveria dispor de tantos clientes quantos fossem os catálogos ou bancos de dados que se quisesse consultar”.

Portanto, o protocolo **Z39.50** é considerado uma ferramenta tecnológica utilizada para facilitar o processo de busca, recuperação e intercâmbio de informações em ambiente digital, e sendo melhor utilizado pelos analistas de sistemas do que pelos bibliotecários.

3.9 **MARC21: uma ferramenta, uma estrutura e um padrão para metadados**

“Diferentemente da **ISO 2709** e do protocolo **Z39.50**, o formato **MARC21** é uma ferramenta de domínio exclusivo do bibliotecário que o utiliza na atividade de descrição bibliográfica” (CÔRTE et.al., 1999, p.248).

Ferramenta é qualquer implemento ou objeto destinado a realizar uma operação ou executar um trabalho de qualquer natureza. Podemos dizer que qualquer que seja o tipo de trabalho realizado, todos nós usamos ferramentas. Por exemplo, o escritor cria com a mente mas necessita de uma ferramenta para anotar seus pensamentos. Um homem que dirige um caminhão, assim como uma pessoa que esteja utilizando um computador está fazendo uso de uma ferramenta, do mesmo modo que aqueles que os fabricaram. O emprego de ferramentas, o desenvolvimento da mecânica, a aplicação da eletricidade, foram os indispensáveis pioneiros de nossos recentes aperfeiçoamentos da automação, da aplicação de controles eletrônicos e das tecnologias de informática (ARNOLD e WHITE, 1965).

As ferramentas das novas tecnologias são as portas de entrada para o mundo globalizado que as unidades de informação de qualidade necessitam para estarem conectadas com essa nova realidade.

A qualidade de uma Unidade de Informação é representada pela sua capacidade em tratar as informações por meio dos processos de armazenamento, recuperação e disseminação com qualidade e padronização e a isso podemos chamar de otimização na gestão da informação. É com essa preocupação que “Bibliotecários e indexadores tem produzido e padronizado metadados por séculos” (MILSTEAD; FELDMAN, 1999, p. 1, tradução nossa), e o Formato de Intercâmbio **Machine Readable Cataloging - MARC21**, ou seja, um registro catalográfico legível por computador é considerado uma estrutura e um

padrão para a construção de metadados produzidos na atualidade. A definição mais comum de metadados é dados sobre dados.

Em recente estudo Grácio (2002, p. 114) define metadados em sua dissertação de mestrado como “um conjunto de elementos que descrevem as informações contidas em um recurso, com o objetivo de possibilitar sua busca e recuperação”. Há também definições de outros autores. Para Souza, Vendrusculo e Melo (2000, p. 93) metadado “significa dado sobre o dado. É a catalogação do dado ou descrição do recurso eletrônico”. E para Souza, Catarino e Santos (1997, p. 2) metadados “são descrições de dados armazenados em banco de dados, ou como é comumente definido dados sobre dados a partir de um dicionário digital de dados”.

Os metadados são um conjunto de dados usados para descrever e representar objetos informacionais, assim como, descrever o seu conteúdo e sua localização e desta forma facilitar o acesso e recuperação de informações seja na Internet ou não, pois o termo metadado se refere a dado em sentido amplo e não somente para fontes eletrônicas. O resumo de um texto ou as suas palavras-chave, por exemplo, representam o texto original, são dados sobre outros dados, portanto considerados metadados.

“Todos os motivos porque a indexação e a catalogação são necessários para fontes impressas se aplicam ainda mais enfaticamente aos metadados para documentos eletrônicos” (MILSTEAD ; FELDMAN, 1999, p. 3, tradução nossa).

De acordo com as autoras Milstead e Feldman (1999) Metadado não é nada de novo para os pesquisadores profissionais, pois eles tem sido capazes de melhorar a precisão de suas buscas através do uso de vocabulários controlados ou limitado as buscas aos descritores, identificadores, autores, títulos, assuntos e outros por muitos anos. Embora metadado tenha se tornado um “buzzword” no mercado de informação, o conceito é importante para, autores e pesquisadores de informação eletrônica. Usado eficientemente, torna a informação acessível

através de rótulos consistentes de seus conteúdos. Entende-se por “buzzword” um termo técnico que soa importante, geralmente usado para impressionar pessoas leigas no assunto.

Para Milstead e Feldman, (1999, p. 1, tradução nossa),

Como o homem que esteve escrevendo prosa toda sua vida sem saber, os bibliotecários e indexadores tem produzido e padronizado metadados há séculos. Ignorando este legado, uma imensa variedade de outros profissionais tem recentemente entrado neste campo, e muitos deles não tem idéia de que alguém mais já “esteve ali e feito isto” antes. Vários sistemas estão sendo desenvolvidos para diferentes e algumas vezes até para o mesmo tipo de informação, resultando numa atmosfera caótica de padrões.

Não importa o nome que se use, catalogação, indexação ou metadados, os profissionais da informação estão familiarizados com os seus conceitos. Segundo Milstead e Feldman (1999, p. 1, tradução nossa),

Seja qual for o nome que se use, catalogação, indexação ou metadados, o conceito é familiar para profissionais de informação. Agora o mundo eletrônico finalmente o descobriu. Até poucos anos atrás, somente alguns filósofos tinham ouvido a palavra “metadado”. Hoje, é difícil encontrar uma publicação que a ignore.

Metadados como um novo conceito surgiu da necessidade de se ter representantes para o conteúdo dos bancos de dados e que pudessem fornecer mais informações sobre eles, devido o crescente número e tipos de objetos digitais disponíveis. De acordo com Milstead e Feldman, (1999, p.3, tradução nossa),

A idéia de metadados como um novo conceito surgiu fora da arena tradicional dos textos e da bibliografia. Com o desenvolvimento dos arquivos de dados, especialmente os dados geoespaciais tornou-se evidente que seria necessário representantes para fornecer mais informações sobre o conteúdo

dos bancos de dados. Um crescente número e tipos de objetos foram disponibilizados digitalmente, mas foi reconhecido que os dados crus tinham pouco valor sem a informação sobre como foi coletado, seu propósito, os formatos, as plataformas para exibição e manipulação dos dados, suas restrições para reprodução e reutilização, assim como as informações de identificação mais convencionais como, autor, título, assunto e resumo.

Diante de tudo isso, os metadados em si não são algo novo, o que é novo hoje em dia são as variedades de padrões que estão sendo desenvolvidos e as formas de utilização (ORTIZ-REPISO JIMÉNEZ, 1999).

Os registros bibliográficos que se tem criado há muitos anos no mundo bibliotecário são essencialmente metadados. Proporcionam informação descritiva e analítica sobre um objeto de informação. [...] os registros **MARC** nos Catálogos de Acesso Público Online - **OPACs**, assim como as fichas catalográficas nos catálogos manuais não são outra coisa senão um conjunto de metadados (ORTIZ-REPISO JIMÉNEZ, 1999, p. 219, tradução nossa).

Para que os metadados alcancem os seus objetivos com sucesso é imprescindível o uso de listas ou bases de dados com assuntos e autoridades controladas, como as linguagens controladas ou tesouros. Por exemplo, se em um catálogo online coletivo não for adotado um padrão para o campo de responsabilidade intelectual, ou seja, autor, permitindo a entrada de várias formas como, Silva, Maria; Silva, Maria A.; ou Maria Aparecida Silva, sem uma padronização para esse campo, caso o usuário queira reunir todas as obras dessa autora, ele provavelmente não conseguirá com as formas variantes como a indicação de autor foi efetuada nesse catálogo, ou seja, não foi usada uma forma autorizada ou padronizada disponível em bases de dados de padronização oficial como a base **BIBLIODATA** ou **Libray of Congress – LC**.

A padronização é fundamental tanto para a representação quanto para a recuperação da informação, sendo necessário a preocupação não somente com os campos para recuperação, mas com a padronização do seu conteúdo.

Na Web os metadados podem ser elaborados no momento da criação de um objeto, como os sites por exemplo ou, então, posteriormente, no caso da tradicional catalogação e indexação dos objetos informacionais, podendo ser confeccionados pelos profissionais bibliotecários no momento do processo da representação descritiva.

No entanto, “na cultura da Web, o autor do documento é a pessoa que aplica o metadado, e como fazer com que milhões de pessoas não profissionais da informação compreendam a importância da catalogação até um certo nível e padrão quando nem sempre os profissionais concordam entre si?” (MILSTEAD e FELDMAN, 1999, p. 5, tradução nossa). Entretanto, numerosos esforços vêm sendo feitos para catalogar recursos da Internet.

A autora Magda El-Sherbini (2000, p. 188, tradução nossa) em seu artigo sobre metadados e o futuro da catalogação faz a seguinte pergunta:

Por que bibliotecas precisam de metadados na era de recursos eletrônicos digitais?

A autora responde com muita pertinência que para lançar esta pergunta, é necessário se lembrar de que os bibliotecários ainda estão lidando com informação que é criada para o mesmo propósito como o livro e o periódico, e então, sujeita à mesma classificação e leis de descrição como todo o resto. E que as bibliotecas usam ferramentas tradicionais para organizar informação que facilita o acesso para materiais impressos e não impressos. AACR2, classificação LC, Dewey, LSCH, e outras ferramentas foram usadas e continuarão sendo usadas para descrever e classificar informação. Provavelmente as bibliotecas continuarão desenvolvendo e melhorando estes antigos padrões assim como elas se adaptaram às demandas das publicações e das indústrias de informação.

“Desde que a informação eletrônica se comporte diferentemente que a mídia de impressão, ela requer tratamento especial para representá-la e classificá-la, e, é aí onde entram os metadados” (EL- SHERBINI, 2000, p. 188, tradução nossa).

A autora ainda transcorre sobre os usos dos metadados e cita algumas sugestões de outras autoras como Milstead e Feldman. Veja a seguir:

Segundo El-Sherbini (2000 p.188, tradução nossa) os metadados são freqüentemente usados para:

- Organizar e manter dados de investimentos das organizações;
- Prover informação para catálogos de dados;
- Prover informação para auxiliar na transferência de dados;
- Auxiliar na descoberta efetiva e na recuperação da informação, permitir aos usuários determinar se o dados encontrados foram realmente os desejados;
- Controlar a informação de acesso restrito e prevenir alguns usuários (como crianças) de ter acesso a esses dados;
- Prover acordo comum como, qual elemento ou qual conteúdo deve ser usado;
- Fornecer informação sobre o uso dos dados como, condições legais de uso, seu tamanho, data e outros;
- Fornecer o histórico dos dados como, a fonte original e qualquer transformação posterior destes dados;
- Fornecer informação sobre o dono ou o criador do texto e
- Fornecer relacionamentos com outros recursos como links para versões atuais e futuras.

De acordo com El-Sherbini (2000 p. 188, tradução nossa) Milstead e Feldman (1999) contam com outros usos para os metadados:

- Garantir que todo o material sobre o mesmo assunto seja encontrado reunido, seja na prateleira ou em um banco de dados online;
- Separar conceitos importantes daqueles que são meramente incidentais no trabalho;
- Assegurar que a mesma informação seja encontrada para cada trabalho, e que ela seja colocada no mesmo lugar, de forma que alguém que esteja procurando os trabalhos por um autor de nome Fields não os encontre misturados com tratados agrícolas sobre fertilização de campos de trigo;
- O uso apropriado de vocabulários de indexação e de estruturas de campos, tanto na pesquisa quanto na catalogação, aumenta a precisão e minimiza a chance de resultados falsos.

Para El-Sherbini (2000, p. 188, tradução nossa), “Os Padrões para Metadados tem um importante papel no suporte ao uso de serviços e recursos eletrônicos. No entanto, para maximizar o uso destes e assegurar consistência, é necessário padrões universalmente aceitados que governem a criação de metadados”.

O Formato MARC é considerado a primeira estrutura e padrão desenvolvidos para a construção e representação de metadados. Possui um conjunto de elementos descritores complexos e rígidos para descrever informações catalográficas e bibliográficas, sendo necessário especialistas como o profissional Bibliotecário para a correta descrição desses elementos (GRÁCIO; 2002; SOUZA; CATARINO; SANTOS, 1997).

A estrutura do registro MARC é deliberadamente complexa, a fim de permitir flexibilidade. Quase todo elemento pode ser usado como ponto de acesso, e cada elemento pode ter qualquer tamanho. Tal complexidade foi projetada numa época quando os catálogos impressos ou em microfimes eram a norma. O formato de registro MARC trouxe uma enorme contribuição à padronização e à comunidade por redes, porém há quem acredite que seja chegado o momento de fazer sua reavaliação. Talvez um formato diferente

seja mais apropriado num ambiente onde predominem os catálogos em linha de acesso público (ROWLEY, 1994, p. 79).

Parece mesmo que este momento chegou fazendo-se necessário uma reavaliação ou mudança no MARC21 considerado inclusive um formato arcaico que requer sistemas especializados para a criação e intercâmbio de seus registros.

Neste cenário surge a emergente Linguagem de Marcação Extendida - XML possibilitando facilidades e interoperabilidades entre sistemas e melhor descrevendo as informações para intercâmbio de dados via Internet. Estaremos abordando a eXtensible Markup Language - XML no próximo capítulo.

CAPÍTULO 4

LINGUAGEM DE MARCAÇÃO XML

A adoção de ferramentas eletrônicas no desempenho das atividades profissionais requer mais mudanças de comportamento do que do aprendizado tecnológico.

Roberto Barssotti

A primeira revolução da informação aconteceu com a Imprensa de Gutenberg. Hoje estamos vivendo a segunda revolução proporcionada pela Internet. Ela realmente está mudando o mundo. A Internet tornou-se essencial para todas as áreas, auxiliando o ser humano nas mais diversificadas atividades, desde uma simples troca de mensagens (e-mails) até operações cirúrgicas monitoradas à distância.

A maior revolução da Internet aconteceu após a criação da World Wide Web em 1991 por Tim Berners-Lee e o desenvolvimento de Browsers como, Mosaic, Netscape, Internet Explorer e outros para possibilitar a visualização das páginas Web.

Atualmente, as páginas e sites da Web tornaram-se parte essencial tanto no mundo empresarial quanto pessoal. Bilhões de páginas estão disponíveis na web e a maioria delas estão escritas em **HyperText Markup Language – HTML**, uma linguagem simples e acessível a qualquer pessoa, usada para a formatação de dados com marcações especiais em um arquivo de texto o qual pode ser visualizado em praticamente qualquer arquitetura computacional (CASTRO, 2001).

Apesar da simplicidade da HTML ter ajudado a alimentar e aumentar a popularidade da Web, sendo possível qualquer pessoa criar uma Homepage, ela possui seus pontos fracos. Apresenta limitações reais em face da imensa massa crescente de informações disponibilizadas a todo momento na Web (CASTRO, 2001). Frente a isso, a **eXtensible**

Markup Language – XML promete revolucionar o mundo de forma mais significativa que a **HTML**.

Baseando-se porém na mesma tecnologia, onde a mãe de todas as linguagens de marcação é a **Standard Generalized Markup Language – SGML**, a **XML** foi especialmente projetada para executar melhor a tarefa de gerenciamento de informações, exigida pelo atual crescimento da Internet.

Para Castro (2001, p.11) “Embora a XML exija um pouco mais de atenção no início, ela fornece muito mais dividendos no final. Resumindo, a HTML permite que todos façam algumas coisas, mas a XML permite que algumas pessoas façam praticamente tudo”.

4.1 O que é XML?

A **eXtensible Markup Language - XML**, ou seja, linguagem de marcação estendida, é descendente da **Standard Generalized Markup Language – SGML**, concebida especialmente para gerenciar, armazenar e transmitir dados via Internet. “Uma linguagem de marcação simplesmente define como você adiciona o significado aos dados, especialmente em um documento” (LIBERTY; KRALEY, 2001, p. 6). Por exemplo: **<título> MARC21 e XML </título>**, essas duas palavras entre tags, marcas ou marcação são o título do documento. A XML “fornece pistas semânticas na forma de tags (marcas) sobre o significado dos dados” (LIBERTY ; KRALEY, 2001, p. 5).

Segundo Ray (2001, p. 2) “uma linguagem de marcação é um conjunto de símbolos que pode ser colocado no texto de um documento para demarcar e rotular as partes desse documento”.

“A Linguagem eXtensible Markup Language (XML) é o resultado do trabalho de um grupo de especialistas estabelecido em 1996 pelo **W3C (World Wide Web Consortium)** ¹, com o objetivo de propor uma simplificação da SGML que fosse voltada às necessidades específicas da Web” (BRYAN, 1998 citado por BAX, 2001, p. 33).

“SGML é uma (meta)linguagem criada há aproximadamente 30 anos”[...] (EDWARDS, 1997 citado por BAX, 2001). Base de um novo paradigma, uma tentativa de padronização dos diversos formatos empregados para representar a informação. Foi reconhecida como um padrão ISO (8879) em 1986. As linguagens de marca ou marcação são aplicadas com vistas a um melhor gerenciamento da informação.

Mas, o autor Erik T. Ray em seu livro intitulado *Aprendendo XML*, defende que “apesar do seu nome XML, não é uma linguagem de marcação, em vez disso, é um kit de ferramentas para a criação, a modelagem e o uso de linguagens de marcação” (RAY, 2001, p. vii)².

Segundo RAY (2001, p.1),

“A Extensible Markup Language (XML) é um kit de ferramentas para armazenamento de dados, um veículo configurável para qualquer tipo de informação e um padrão aberto e em evolução, abraçado por todos, desde banqueiros até webmasters. Em apenas alguns anos, ela capturou a imaginação tanto de mestres da tecnologia quanto de especialistas do setor.”

E, um outro erro de conceito, é de que a XML substituirá a HTML, o autor esclarece dizendo:

Na verdade, a HTML está sendo absorvida gradativamente pela XML, tornando-se uma versão mais clara de si mesma, chamada XHTML. E isso é

¹ W3C – Organização que se encarrega do desenvolvimento e manutenção dos padrões da Web. Maiores informações visite o site www.w3c.org

² Citação retirada do prefácio do livro.

apenas o começo, pois a XML permitirá a criação de centenas de novas linguagens de marcação, abordando cada tipo de aplicação e documento (RAY, 2001, p. vii)³.

Ray (2001, p.1) atribui o segredo do sucesso da XML a uma pequena lista de suas características:

- A XML pode armazenar e organizar praticamente qualquer tipo de informação em um formato adequado às suas necessidades.
- Como um padrão aberto, a XML não está ligada às fortunas de qualquer empresa isolada, e nem acoplada a qualquer software em particular.
- Com o Unicode como seu conjunto de caracteres padrão, a XML aceita um número incrível de sistemas escritos (scripts) e símbolos, desde caracteres rúnicos escandinavos até ideógrafos han chineses.
- A XML oferece muitas maneiras de verificar a qualidade de um documento, com regras para sintaxe, verificação de vínculo interno, comparação com modelos de documento e tipos de dados.
- Com sua sintaxe clara e simples e sua estrutura sem ambigüidades, a XML é fácil de ler e analisar por seres humanos e programas.
- A XML é facilmente combinada com folhas de estilo para criar documentos formatados em qualquer estilo que você desejar. A pureza da estrutura da informação não atrapalha as conversões de formato.

A XML é a promessa de facilitar o acesso ao grande volume de informação que está dentro do nosso alcance, mas as limitações da tecnologia existente podem dificultar o acesso (RAY, 2001).

Mas afinal, o que é XML?

Para Ray (2001, p.2) esta é uma pergunta difícil de responder, pois

³ Citação retirada do prefácio do livro.

Em um nível, XML é um protocolo para conter e gerenciar informações. Em outro, é uma família de tecnologias que pode fazer tudo, desde formatar documentos até filtrar dados. E, no nível mais alto, é uma filosofia para o tratamento de informações, que busca o máximo de utilidade e flexibilidade para os dados, refinando-os à sua forma mais pura e mais estruturada. Um conhecimento completo da XML focaliza todos esses níveis.

Somente para lembrar, visto que já foi citado no capítulo anterior,

Protocolo é um conjunto de regras e comandos pré-estabelecidos entre os componentes de uma rede para que se efetue suas comunicações, ou seja, para que os computadores envolvidos “falem” entre si. Os protocolos são, geralmente, documentos em páginas de material impresso, porém, como ocorre com tantas outras tecnologias complexas, o usuário não precisa conhecê-los em detalhes para que possa deles se utilizar (KROL, 1993 citado por FERREIRA, 1994, p. 258).

A linguagem XML ainda é recente, foi criada em 1996 e liberada oficialmente pelo World Wide Web Consortium – W3C (órgão responsável pelo seu desenvolvimento e manutenção) em 1998, sendo, portanto, a sua introdução na Web no final da década de 1990, por isso ainda é muito discutida e analisada por especialistas no assunto. Liberty e Kraley (2001, p.5) afirmam, “XML, uma das mais novas e sem dúvida, uma das mais poderosas tecnologias da Web”.

Para Bax (2001, p. 32),

Após a surpreendente evolução da microinformática nas últimas duas décadas, que elevou sobre maneira o nível de abstração da relação homem/máquina, constata-se hoje que a maior contribuição visível dos computadores está em sua utilização como ferramentas de armazenamento, organização, recuperação e intercâmbio de informações entre pessoas, empresas e até entre programas.

Enfim, hoje o computador é visto pela sociedade, cada vez mais, como uma ferramenta de comunicação, e não propriamente de cálculo.

4.2 Origem da XML

Com o dilúvio de informação que o século XX proporcionou a humanidade, considerada uma era de informação sem paralelos na história humana, hoje temos tantos dados sobre o universo que nem sabemos o que fazer com todas as informações, deixando claro que a organização de todo esse conhecimento é uma questão fundamental.

Os primeiros formatos eletrônicos que surgiram se preocupavam mais em descrever como as informações seriam apresentadas que com a estrutura e o significado do documento. Por exemplo: as linguagens de formatação Troff e TEX fizeram um trabalho fantástico para formatar documentos impressos, mas perderam qualquer senso de estrutura. Os documentos eram limitados a serem vistos na tela ou impressos em papel, não era possível escrever programas facilmente para pesquisar e canalizar informações, fazer referência cruzada eletronicamente ou redefinir a finalidade dos documentos para aplicações diferentes (RAY, 2001).

O problema foi resolvido com a codificação genérica que usa tags descritivas em vez de código de formatação. A **Graphic Communications Association – GCA** foi a primeira organização a explorar essa idéia. E no final da década de 1960 diferentes tipos de documentos estavam sendo codificados com tags genéricas (RAY, 2001).

Um grande avanço iniciou nos primórdios dos anos 80 quando a **IBM** começou a trabalhar com a Organização Internacional para Padronização - **ISO** para produzirem um documento padrão de linguagem de marcação baseado no seu produto **Generalized Markup Language - GML**. O resultado desse processo de padronização foi a **Standard Generalized Markup Language – SGML**, ou seja, Linguagem de Marcação Generalizada Padrão (FIANDER, 2001).

A SGML não é diretamente um padrão de formatação de texto. Ela define uma “linguagem” especializada para se descrever a estrutura de documentos, ou qualquer tipo de dado textual, deixando a interpretação do dado para

outro programa. Dessa forma a SGML é ideal para codificar dados textuais fortemente estruturados para comunicação entre sistemas de computador (FIANDER, 2001, p.18, tradução nossa).

De acordo com Fiander (2001, p. 18, tradução nossa),

Infelizmente, a SGML é um amplo e complicado padrão, os recursos de computação necessários para suportar o padrão completo, com todas as suas opções, não pode ser amplamente empregado atualmente. Essa dificuldade levou ao desenvolvimento da eXtensible Markup Language – XML pelo Consórcio da World Wide Web – W3C, a qual é um subproduto da Standard Generalized Markup Language - SGML que foi desenvolvido para facilidade de implementação e para interoperabilidade com ambas, SGML e HTML.

É importante entender que a SGML é uma meta(linguagem) usada para marcação de textos em qualquer formato e não uma linguagem para marcação de informação usada para a Internet. De acordo com Furgeri (2001, p. 41),

Em 1996, especialistas em SGML, a principal linguagem de marcação da qual surgiu a HTML, sob a chefia de Jon Bosak, da Sun Microsystems, se uniram para definição de **um novo padrão de marcação que pudesse ser utilizado na Internet**, constituindo-se em uma versão simplificada da SGML [...]. No final de 1996, o comitê de trabalho anunciou a primeira versão preliminar da XML. (grifo nosso).

Segundo Ray (2001, p.11),

[...] algumas pessoas tentaram adaptar a SGML para a Web – ou melhor, adaptar a Web à SGML. Isso provou ser muito difícil. A SGML era muito grande para ser colocada em um browser da Web pequeno. Era necessária uma linguagem menor, que ainda mantivesse a generalidade da SGML – assim nasceu a Extensible Markup Language (XML).

4.3 Vantagens da XML

A XML está se tornando uma ferramenta virtualmente indispensável para todo tipo de conteúdo que é digitalmente produzido e portanto uma poderosa ferramenta que promete muitas mudanças em todas as áreas.

Para Rhyno (2002, p. 97, tradução nossa) em um nível geral, XML tem pelo menos três vantagens chave:

1 - XML é padrão. Um subconjunto da Linguagem de Marcação Generalizada Padrão (SGML), um padrão ISO bem estabelecido e altamente respeitado. XML também oferece mecanismos (definições de tipo de documento, e mais recentemente, XML Schema) para encorajar e obrigar definições comuns da estrutura lógica de um documento, assim os próprios documentos padrões são mais fáceis de aplicar.

2 - XML é extensível. Não tem nenhum limite no número de linguagens de marcação que pode ser definido usando-a, e cada uma destas linguagens pode consistir em números ilimitados de etiquetas. Esta flexibilidade é essencial para favorecer a difundida adoção da linguagem e ajuda a explicar por que XML está sendo utilizada para linguagens de marcação em todas as áreas concebíveis de estudo, percorrendo desde jogo de xadrez até a leitura da mente.

3 - XML é bem suportada. Parece que agora é esperado que toda tecnologia séria da Web defina suas relações para XML, a qual favoreceu uma explosão de tecnologias relacionadas. O W3C tem feito muitas ferramentas e kits de ferramentas livremente disponível para desenvolvedores da Web, e utilidades e aplicações para XML estão sendo desenvolvidas por uma comunidade mundial de vendedores, contribuintes de fontes abertas, e outros grupos que buscam mover informação pela Web.

Segundo Furgeri (1999, p. 43), as principais vantagens oferecidas pela utilização da XML são:

- Os browsers, apoiados por linguagens de programação como Java ou JavaScript, podem fazer grande parte do trabalho de processamento; o conteúdo do documento pode ser manipulado e reorganizado, cálculos podem ser realizados para gerar novos conteúdos instantaneamente, o que proporciona a geração de outros documentos.
- Um documento, com mesmo conteúdo, pode ser visualizado de diferentes formas, para diferentes usuários, através da utilização de folhas de estilo.
- O conteúdo de um documento pode ser pesquisado de maneira inteligente, com base em sua estrutura (por exemplo, um browser pode ser orientado a encontrar quais dos livros, presentes no documento, possuem um determinado autor).
- As ferramentas de criação, baseadas em XML, permitirão que os autores concentrem-se no conteúdo do documento e não na sua formatação.
- Permite um alto grau de automação para agentes de software.
- Os documentos XML são essencialmente banco de dados de informações, podendo ser consultados e processados como qualquer banco de dados tradicional.
- Possibilita a criação de Tags conforme a necessidade.
- A XML possibilita a entrega de qualquer tipo de informação estruturada por toda a WEB. Aplicações XML padronizadas possibilitarão que diferentes aplicativos trabalhem em conjunto, significando uma maior interoperabilidade.
- Do ponto de vista de um publicador, a informação que ele mantém é o seu "Capital de Giro". A XML permite que esse bem seja administrado de uma maneira coerente e eficiente. Grandes recursos de informação podem ser divididos em micro documentos, os quais podem ser mantidos e administrados usando técnicas de banco de dados.
- Os recursos de circulação da XML podem ser usados para criar uma rede WEB de conhecimento (knowledge WEB), ou até várias redes superpostas que, de outra forma, seria uma massa de documentos sem conexão. Essas

circulações podem ser mantidas e preservadas, separadas dos próprios documentos, simplificando a tarefa de manutenção dos links.

- Do ponto de vista do cliente, receber XML em vez de HTML, torna-o muito mais auto-suficiente. Quando a informação codificada em XML está relacionada a uma aplicação específica de interesse do cliente, seu valor aumenta dramaticamente. Quando a informação é guardada em uma aplicação XML específica, sua marcação pode refletir precisamente as semânticas daquela informação.

Essas são algumas das vantagens principais que a XML oferece, e que facilitam algumas tarefas como, o tratamento, o armazenamento, a recuperação e intercâmbio de informações entre muitas outras.

4.4 Sintaxe básica da XML

A XML permite marcar os documentos levando em conta a semântica da informação, ou seja, cada pessoa ou instituição pode criar um conjunto de etiquetas que melhor represente os elementos que compõem os seus documentos (MARTÍNEZ GONZALEZ, 2000).

A esta propriedade que possibilita a criação de novas etiquetas é denominada extensibilidade. Desta propriedade se deriva uma outra, a legibilidade. Visto que, um documento XML etiquetado com marcas suficientemente representativas, qualquer usuário, com ou sem conhecimentos em XML é capaz de reconhecer os distintos elementos de um documento XML [...] (MARTÍNEZ GONZALEZ, 2000 p, 125, tradução nossa).

Observe os dois exemplos de documento XML:

```
<?xml version="1.0" ?>
<!-- Exemplo de um documento XML -->
- <Cadastro>
  <nome>Sofia</nome>
  <idade>25</idade>
  <email>sofiaet@tvnet.com</email>
</Cadastro>
```

doc. 1

```
<?xml version="1.0" ?>
<!-- Exemplo de um documento XML -->
- <Catalogo>
  <Autor>Flamino, Adriana Nascimento</Autor>
  - <Titulo>
    MARC21 e XML como ferramentas para a
    consolidação da Catalogação Cooperativa
    Automatizada
    <subtitulo>uma revisão de literatura</subtitulo>
  </Titulo>
  <Ano>2003</Ano>
</Catalogo>
```

doc 2

FIGURA 23: Exemplos de documentos em XML

A expressão `<cadastro>` é chamada de marca de início e `</cadastro>` de marca de fim, assim como as outras tags como `catalogo`, `nome`, `autor` e outras. Pode-se observar que todas elas tem a sua marca de início e fim. As marcas de início e de fim são também chamadas de marcações (ABITEBOUL; BUNEMAN; SUCIU, 2000).

As marcas ou marcações em XML de acordo com Abiteboul, Buneman e Suciú (2000, p. 28),

são definidas pelos próprios usuários, não há marcas predefinidas como em HTML. O texto entre uma marca de início e a correspondente marca de fim, incluindo as marcas embutidas, é chamado um elemento e as estruturas entre as marcas são o conteúdo. O termo subelemento também é utilizado para descrever a relação entre um elemento e seus elementos componentes.

Assim, `<Titulo> ... </Titulo>` é um subelemento de `<catalogo> ... </catalogo>` no exemplo citado (FIG. 23).

Veja o exemplo do documento (doc.1 – figura 23) em XML, agora em HTML:

```
<HTML>
<Body>
<h1> Cadastro </h1>
<p> <b> Sofia </b>, 25, <i> Sofiaet@tvnet.com </i> </p>
</Body>
</HTML>
```

FIGURA 24 : Exemplo de um documento em HTML

A informação no documento HTML na figura 24 é essencialmente a mesma que a do documento XML doc.1 na figura 23. Ambos descrevem o cadastro de uma pessoa de nome Sofia, 25 anos e seu e-mail. Mas, enquanto a HTML descreve a apresentação, a XML descreve o conteúdo. Uma aplicação pode facilmente entender os dados XML (ex., nome separado de idade e de e-mail); por outro lado, não há indicação em XML de como os dados devem ser exibidos (ABITEBOUL; BUNEMAN; SUCIU, 2000).

A XML se utiliza de tecnologias associadas a ela para a apresentação dos seus dados como, a eXtensible Stylesheet Language – XSL , ou seja, linguagem de folha de estilos, que veremos adiante.

Veja um outro exemplo de um mesmo documento em HTML e em XML do autor Furgeri (2001, p. 45), em seu exemplo ele supôs ser necessário a criação de um documento HTML para dar informações para venda de microcomputadores. As mesmas informações serão colocadas nos dois documentos, HTML e XML, para que desta forma fique mais fácil de perceber a principal diferença entre as duas linguagens, que se refere ao significado do texto. Observe os exemplos:

Em HTML:

```
<HTML>
<BODY>
Microcomputador Pentium IV, 1.5 GHz, 256MB de RAM, Monitor 17 polegadas,
Mouse, Teclado, estabilizador.
</BODY>
</HTML>
```

FIGURA 25: Documento em HTML

FONTE: FURGERI, 2001, p. 145

Em XML:

```
<?xml version="1.0"?>
<MICROCOMPUTADOR>
  <MODELO> Pentium IV </MODELO>
  <VELOCIDADE> 1,5 GHz </VELOCIDADE>
  <RAM> 256MB </RAM>
  <MONITOR> 17 Polegadas </MONITOR>
  <TECLADO> Sim </TECLADO>
  <MOUSE> Sim </MOUSE>
  <ESTABILIZADOR> Sim </ESTABILIZADOR>
  <IMPRESSORA> Não </IMPRESSORA>
</MICROCOMPUTADOR>
```

FIGURA 26: Documento em XML

FONTE: FURGERI, 2001, p. 145

Ao comparar os dois documentos, fica visível de que a linguagem XML é bem parecida com a HTML, possuindo também tags para marcar as informações, no entanto existe uma diferença fundamental entre elas: a XML torna o documento mais “inteligente”, criando uma estrutura que demonstra claramente qual o significado das informações, enquanto a HTML se preocupa com a apresentação do conteúdo e não com o seu significado (FURGERI, 2001).

Apesar de os dois documentos conterem as mesmas informações, no documento escrito em HTML um software não pode reconhecer o significado dos dados inseridos, pois estão misturados em um texto corrido.

Isso já não acontece com o documento escrito em XML, pois as informações estão separadas e organizadas. Assim como um software, o leitor pode ler o conteúdo do documento e interpretar corretamente seus significados [...] (FURGERI, 2001, p. 46).

Segundo Gazan (2000, p. 14, tradução nossa),

A XML não é a HTML com novas tags. As tags ou etiquetas em um documento HTML são fixas, e faz pouco mais que instruir um browser como exibir o conteúdo. Elas não provêem nenhuma ajuda nem posições seguras para as máquinas de busca que tentam ter acesso ao conteúdo de um documento. Por outro lado, as etiquetas XML descrevem o conteúdo do documento e expressam as relações hierárquicas entre elementos de dados, muito semelhante aos campos dos registros em um banco de dados.

As duas linguagens de marcação HTML e XML, apesar de serem bastante parecidas existem diferenças significativas entre elas. De acordo com Martínez González (2000); Castro (2001); Ramalho, (2002):

- A XML tem como finalidade descrever o conteúdo dos dados enquanto a HTML se preocupa com a apresentação dos mesmos;
- A XML possui extensibilidade, as suas etiquetas são definidas de acordo com a necessidade de seus usuários, enquanto a HTML possui etiquetas predefinidas;
- A XML possui links mais potentes do que a HTML, permitindo recursos que esta não possui;
- A XML permite o armazenamento da informação e sua descrição em um mesmo local;
- A XML foi criada especialmente para a recuperação e o intercâmbio de dados;

- A sintaxe rígida da XML facilita o trabalho das aplicações informáticas, enquanto a sintaxe da HTML dificulta o tratamento dos seus dados pelos browsers, justamente pela falta de rigidez.

Por exemplo, um documento HTML pode ter etiqueta de abertura e não possuir a etiqueta de encerramento. (um caso muito habitual são as páginas que contém uma etiqueta <html> mas não possuem em nenhum lugar do documento a etiqueta </html>). Esta diferença faz com que o tratamento automático dos documentos HTML seja muito mais difícil [...]. Isso não acontece com os documentos XML [...], uma vez que qualquer documento que se diz XML deve seguir regras sintáticas e ser bem formado (MARTÍNEZ GONZÁLEZ, 2000, p. 126, tradução nossa).

Um documento XML bem formado é aquele que possui tags corretas, ou seja, para toda tag de abertura existe sua correspondente de encerramento, assim como a padronização nos tamanhos das letras, sendo maiúsculas ou minúsculas. Possui uma tag raiz do documento na qual todas as outras tags e dados estão cuidadosamente aninhados, tornando possível um software XML interpretá-lo como uma estrutura hierárquica (FURGERI, 2001).

Porém, um documento XML bem formado, não significa que ele seja válido. Existem processadores XML que validam ou não o documento, verificando se todas as tags usadas no documento estão de acordo com os padrões previamente definidos pela **Document Type Definition – DTD**, ou seja, definição de tipo de documento, que mantém as características de cada uma das tags como, conteúdo, obrigatoriedade, seqüência e outros. Para um documento XML ser válido, obrigatoriamente ele deve ser bem formado. Um dos processadores XML mais simples que realizam a validação de documentos XML chama-se XMLValidator (FURGERI, 2001; RAMALHO, 2002).

A **HyperText Markup Language – HTML**, segundo Castro (2001, p. 12) deve o seu sucesso “à sua simplicidade, facilidade de uso e tolerância”.

A HTML aceita muitas coisas: não se importa se as letras são maiúsculas ou minúsculas, é flexível quanto ao uso de aspas, não se preocupa muito com as tags de fechamento. Sua tolerância a torna acessível a qualquer pessoa (CASTRO, 2001, p. 12).

A falta de rigidez e padronização da HTML quanto ao uso de maiúsculas, minúsculas e pontuação, dificulta o trabalho dos navegadores para exibir corretamente o conteúdo de uma página em HTML (CASTRO, 2001). No entanto, essa falta de rigidez permite que qualquer pessoa possa utilizá-la, devido a sua simplicidade.

“Mas a simplicidade da HTML limita seu poder. Como as tags da HTML são principalmente para formatação, elas não fornecem informações sobre o conteúdo de uma página da Web e assim fica difícil reutilizar essas informações em outro contexto” (CASTRO, 2001, p. 12).

Para Castro (2001) a XML é a resposta à tolerante, porém limitada HTML. Aparentemente a XML parece muito com a HTML, tags, atributos e valores. Porém, a XML vai mais longe do que a HTML,

[...] em vez de funcionar como uma linguagem apenas para criar páginas na Web, a XML é uma linguagem para criar outras linguagens. Você utiliza a XML para criar sua própria linguagem de marcações personalizadas e depois utiliza essa linguagem para formatar seus documentos (CASTRO, 2001, p. 13).

Para Castro (2001, p. 13) o poder da XML está nos seguintes detalhes:

Se uma tag identifica dados, os dados ficam disponíveis para outras tarefas. É possível escrever um programa para extrair apenas as informações necessárias, unindo-as talvez com dados de outra origem e finalmente fornecendo a combinação resultante em outra forma, para outra finalidade.

Em vez de ficarem perdidas em uma página da Web baseada em HTML, as informações rotuladas podem ser reutilizadas sempre que necessário.

Porém, segundo Castro (2001, p.13),

[...] como sempre, o poder tem seu preço. XML não tem nem um pouco da tolerância da HTML. A fim de facilitar para os parsers de XML – programas que lêem e interpretam dados em XML, seja de forma independente, seja dentro de um navegador, a XML exige uma atenção cuidadosa ao uso de maiúsculas e minúsculas, aspas, tags de fechamento e outras minúcias ignoradas pelos autores de HTML.

O autor porém, acredita que esse caráter meticuloso da XML possa impedi-la de se tornar uma ferramenta para a criação de páginas pessoais na Web.

No entanto, a XML com certeza oferece aos profissionais da informação o poder de gerenciar e intercambiar informações em grande escala (CASTRO, 2001).

4.4.1 Ferramentas para a criação de documentos XML

Quanto as ferramentas para criar documentos em XML, pode ser utilizado qualquer editor de texto ou processador de palavras, como o Notepad ou o Wordpad no Windows ou TeachText ou SimpleText no Macintosh. Por ser uma linguagem de código aberto assim como a HTML, podem utilizar diversos programas desenvolvidos por terceiros para facilitar a criação de arquivos e estruturas de dados (CASTRO, 2001; RAMALHO, 2002).

Muitos editores de HTML foram desenvolvidos para facilitar a criação de páginas HTML, mas atualmente as atenções das indústrias estão voltadas à criação de editores XML. Fazendo uma busca na Internet pode-se ter acesso e permissão para download de vários

editores XML como, XMLSPY, MSXML, PETER'S XML EDITOR entre outros (RAMALHO, 2002). Veja a tela do editor XMLSHELL (FIG. 27):

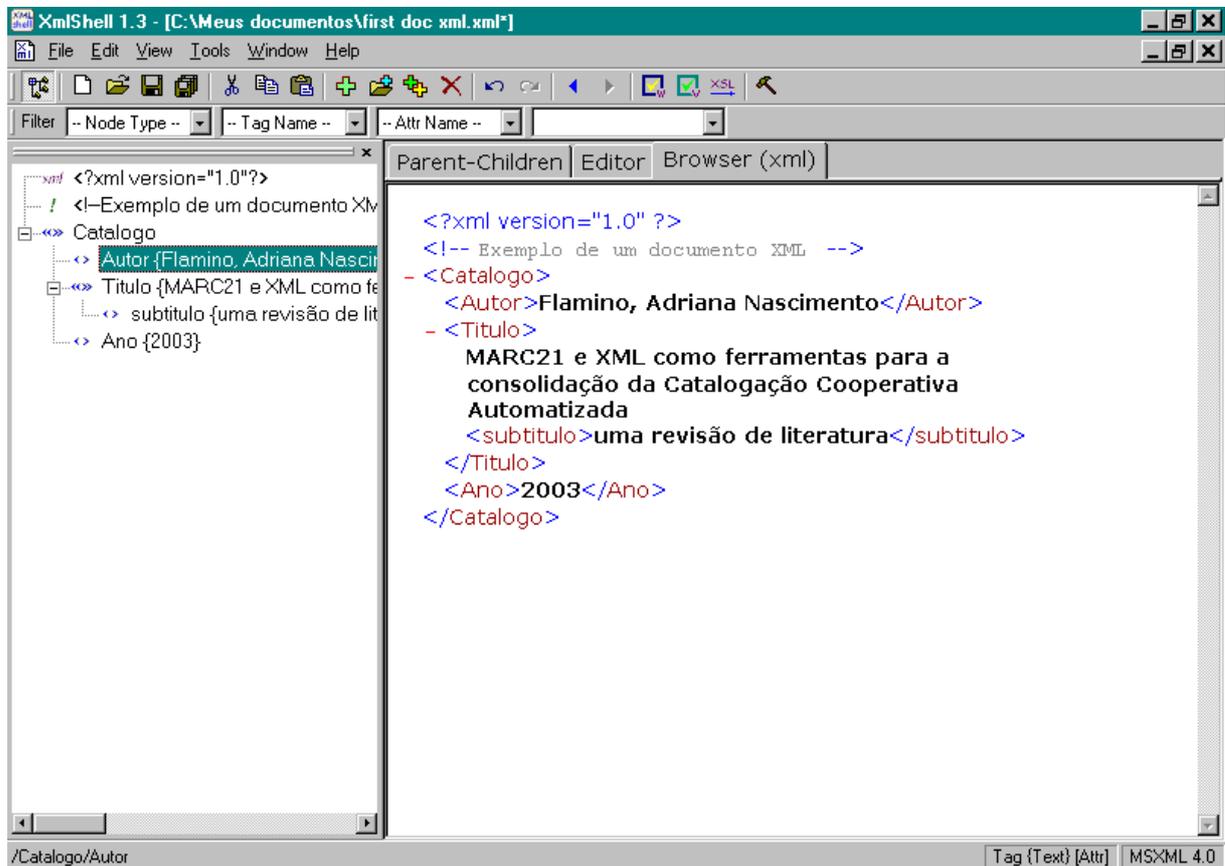


FIGURA 27: Tela do editor XMLSHELL

4.5 TECNOLOGIAS ASSOCIADAS À XML

Para Castro (2001) a XML em si é muito simples. São as tecnologias associadas a ela que a tornam poderosa. São elas: DTD, XML Schema, XSL, Xlink, Xpointer, Xpath, XML Namespaces entre outras.

4.5.1 Document Type Definition - DTD

A característica principal da XML é a facilidade que ela proporciona para modelarmos os dados. Por meio da modelagem dos dados, pode-se especificar quais elementos o documento terá, assim como a ligação dos elementos entre si. A modelagem dos dados pode ser feita através da Document Type Definition – DTD, ou seja, declaração de tipo de documento ou por intermédio de esquemas XML – XML Schemas (RAMALHO, 2002).

“A estrutura dos dados pode ser chamada de esquema ou até mesmo de vocabulário. Ao definir um esquema, pode-se criar regras para que os dados de um documento obedeçam a determinadas regras pré-estabelecidas” (RAMALHO, 2002, p. 31).

A DTD define a estrutura do documento, assim como especifica uma lista dos elementos válidos e seu possível conteúdo. Quando se trabalha sozinho, a pessoa ou instituição tem a liberdade de definir seus elementos do jeito que quiser. No entanto, como a XML é uma linguagem cuja principal finalidade é compartilhar dados, deve-se pensar nas outras pessoas ou outras instituições que estejam envolvidas no processo de criação e intercâmbio dos dados (RAMALHO, 2002).

Ao utilizar a DTD pode-se criar um modelo de dados que pode ser usado por outros grupos de pessoas ou instituições envolvidas no processo de intercâmbio de dados e por meio dela verificar se os dados recebidos de terceiros são válidos e se estão seguindo as normas definidas (RAMALHO, 2002).

Segundo Ramalho (2002, p. 32) “a definição do tipo de documento pode ser interna ou externa ao documento XML.”

- Uma DTD externa pode ser usada por diversos documentos.
- Um documento que use uma DTD interna é completamente independente de fontes externas.
- Nos dois casos, o formato da DTD é o mesmo (RAMALHO, 2002, p. 32).

➤ DTD INTERNA

Uma DTD interna é criada dentro do próprio arquivo XML e não depende de arquivos externos. A tag **DOCTYPE** é usada para declarar uma DTD.

Veja um exemplo de sua sintaxe (FIG. 28):

```
<?xml version= "1.0" encoding= "ISO8859-1" ?>
<!DOCTYPE relatório [
    <!ELEMENT relatório (para, de, título, mensagem)>
    <!ELEMENT para (#PCDATA)*>
    <!ELEMENT de (#PCDATA)*>
    <!ELEMENT título (#PCDATA)*>
    <!ELEMENT mensagem (#PCDATA)*>
]>
<relatório>
<para>FAPESP</para>
<de>Flamino, Adriana Nascimento</de>
<título>MARC21 e XML</título>
<mensagem>Relatório de início; o científico; fica</mensagem>
```

FIGURA 28: Exemplo de uma DTD interna

Analisando o exemplo:

!ELEMENT relatório define que o elemento relatório possui quatro elementos (para, de, título e mensagem).

!ELEMENT para define que o elemento para é do tipo “CDATA”, assim como as demais linhas. “CDATA refere-se a dados do tipo caractere, porém com uma diferença: o texto contido em CDATA não será processado pelo parser, ou seja, referências a entidades e comandos de marcação não serão considerados” (RAMALHO, 2002, p. 33).

Na tag <mensagem> estão inseridos alguns caracteres unicode que estão representando as letras ó, çã e í. “Unicode é um padrão para a representação de caracteres

pertencentes a qualquer idioma. Mais de 65 mil caracteres podem ser associados com um número hexadecimal equivalente ao valor de 16 bits do caractere” (RAMALHO, 2002, p.27).

➔ DTD EXTERNA

Uma DTD externa é criada independentemente do arquivo XML, ou seja, além do arquivo XML é criado um arquivo de extensão **.dtd**. Enquanto a DTD interna é criada dentro do arquivo XML, a DTD externa é criada fora do arquivo XML e pode ser usada por diversos documentos, o que não ocorre com a DTD interna. Mas o formato de uma DTD interna e externa é o mesmo.

Veja o exemplo:

```
<!DOCTYPE relatório SYSTEM "relatório.dtd">
  <para>FAPESP</para>
  <de>Flamino, Adriana</de>
  <titulo>MARC21 e XML</titulo>
  <mensagem>Relatório de iniciação científica</mensagem>
```

FIGURA 29: Exemplo de uso de uma DTD externa

O conteúdo do arquivo **relatório.dtd** referenciado pela instrução DOCTYPE é o mesmo da DTD interna. Observe:

```
<?xml version="1.0" encoding="ISO8859-1" ?>
<!ELEMENT relatório (para, de, título, mensagem)>
<!ELEMENT para (#PCDATA)*>
<!ELEMENT de (#PCDATA)*>
<!ELEMENT título (#PCDATA)*>
<!ELEMENT mensagem (#PCDATA)*>
```

FIGURA 30: Exemplo de um arquivo DTD

De acordo com Ramalho (2002, p. 34) “o uso da DTD interna ou externa não afeta a aparência dos dados”.

Segundo Ray (2001) a Document Type Definition – DTD, é o tipo mais popular de modelo de documento. “As DTDs, na realidade são mais antigas que a XML, sendo um prosseguimento da SGML, com a sintaxe quase completamente intacta” (RAY, 2001, p. 153).

Para Ray (2001, p. 53) “uma DTD define um tipo de documento da seguinte maneira”:

- Declara um conjunto de elementos permitidos. Não se pode usar qualquer nome de elemento além daqueles nesse conjunto. Funciona como o “vocabulário” da linguagem.
- Define um modelo de conteúdo para cada elemento. O modelo de conteúdo é um padrão que diz quais elementos ou dados podem entrar em um elemento, em que ordem, em que quantidade e se eles são obrigatórios ou opcionais. Funciona como a “gramática” da linguagem.
- Declara um conjunto de atributos permitidos para cada elemento. Cada declaração de atributo define o nome, o tipo de dados, valores default (se houver) e comportamento (por exemplo, se é obrigatório ou opcional) do atributo.
- Oferece uma série de mecanismos para facilitar o gerenciador do modelo, por exemplo, o uso de entidades de parâmetros e a capacidade de importar partes do modelo a partir de um arquivo externo.

Segundo Martínez Gonzalez (2000), cada usuário ou comunidade de usuários pode definir suas próprias DTD adaptadas ao tipo de informação que manipula. Por exemplo, A **Text Encoding Initiative - TEI** é uma grande DTD concebida para representar documentos de caráter geral. Enquanto a Docbook é uma DTD concebida para ser utilizada por documentos de caráter técnico.

As DTDs em seu estado atual, apesar do seu valor, sua capacidade expressiva é limitada, não se pode restringir possíveis valores de um elemento ou atributo, por exemplo, não é possível indicar que um preço deve ser sempre maior que 0 (zero). Por causa dessa limitação e de outras que já foram detectadas nas DTDs (MARTÍNEZ GONZALEZ, 2000), ela está sendo substituída por um novo sistema com o intuito de resolver esses problemas. Segundo Castro (2001, p.35) as DTDs são consideradas “um sistema antiquado, mas amplamente utilizado”.

O novo sistema de ponta que vem substituir as DTDs é o XML Schema desenvolvido e recomendado pelo W3C.

4.5.2 XML Schema

O esquema XML - expressa vocabulários compartilhados e permite que máquinas executem regras elaboradas por pessoas. Eles provêm meios para definir a estrutura, conteúdo e semântica de documentos XML. O uso deste esquema é recomendado pelo World Wide Web Consortium – W3C (W3C, 2000-2003).

Esquemas são conjuntos de regras que especificam quais elementos e atributos são permitidos ou necessários em um documento (CASTRO, 2001).

Segundo Castro (2001, p. 69) “as DTDs têm diversas desvantagens em relação aos esquemas escritos com XML Schema”:

Primeiro, as DTDs são escritas em uma sintaxe que tem pouca relação com XML e que não podem ser analisadas com um parser XML. Em segundo lugar, todas as declarações em uma DTD são globais, o que significa que você não pode definir dois elementos diferentes com o mesmo nome, mesmo se aparecerem em contextos separados.

Finalmente, e talvez o mais importante, as DTDs não podem controlar que tipo de informação determinado elemento ou atributo podem conter.

Segundo o autor, o esquema XML tenta solucionar cada um desses problemas e ainda permite muito mais controle sobre o conteúdo de um documento XML.

De acordo com Ioannides (2000, p. 12, tradução nossa) o novo XML schema promete ser:

- mais expressivo que XML DTDs;
- expressado em XML;
- auto-descritivo;
- utilizável por uma larga variedade de aplicações que empregam XML;
- utilizável diretamente na Internet;
- aperfeiçoado para interoperabilidade;
- simples o suficiente para ser implementado em projetos modesto;
- coordenado pelo W3C.

Para o autor as diferenças entre a DTD tradicional e o novo XML Schema são facilmente percebidas por esses dois itens:

- (1) os novos esquemas XML preservam a estrutura lógica da DTD mas não o físico.
- (2) os novos esquemas XML possuem mais elementos que as DTDs, mas eles são muito mais fáceis de ler.

Para Ioannides (2000, p. 14, tradução nossa),

O novo XML schema não é somente uma tentativa para simplificar esquemas já existentes, mas um esforço para criar uma linguagem capaz de definir conjuntos de regras para qualquer e possível recursos de dados[...]. É de extrema importância ter uma linguagem para esquemas que seja universalmente aceita, pois sem isto, nenhuma migração séria de dados para

a estrutura XML será possível. Bases de dados com estruturas de dados de difícil controle como MARC serão grandemente beneficiadas com tal migração. Não dependemos mais de estruturas de dados que foram pré-definidas a décadas atrás para satisfazer diferentes necessidades.

A flexibilidade e a extensibilidade da XML para a criação de estruturas de dados é extremamente valiosa, devido às diversas necessidades atuais de informações. Segundo Ioannides (2000, p. 14, tradução nossa),

A natureza extensível desta linguagem para esquemas permitirá a fácil criação de qualquer estrutura de dados, provendo assim uma flexibilidade exigida pela mutabilidade de necessidades informacionais nos dias de hoje. Trazendo uma riqueza de metadados em um formato extensível e permitindo tirar proveito da natureza dinâmica do trabalho em rede, esses são aspectos extremamente excitantes para os profissionais da informação. O primeiro passo, prematuro talvez, contudo muito simbólico, foi a criação de DTD para o formato MARC pelo projeto CORC. No futuro, esquemas inteligentes permitirão misturar os metadados existentes com texto completo, multimídia, e muito mais. As possibilidades são infinitas.

4.5.3 Folhas de estilos – CSS e XSL

“As folhas de estilos são conjuntos de regras que se adicionam aos documentos XML e [HTML] e que se aplicam a estes no momento de sua apresentação ao usuário” (MARTÍNEZ GONZALEZ, 2000, p. 132, tradução nossa).

As folhas de estilos mais conhecidas são: **Cascading Style Sheets – CSS** e **eXtensible Stylesheet Language – XSL**.

A CSS especifica ao navegador ou browser as características da apresentação do documento ao usuário como, o tipo de cor, tamanho da letra, o tipo de fonte, espaçamento e outros. Com o uso da CSS podemos, segundo Ramalho (2002, p. 15),

[...] separar a formatação aplicada aos elementos dos arquivos que contêm os elementos. Podemos criar diversas folhas de estilo e aplicá-las a um determinado elemento, de forma a obtermos diferentes visualizações dependendo da conveniência. Por exemplo, uma página que fale sobre futebol pode exibir um mesmo texto nas cores verde e branco para o usuário palmeirense, ou preto e vermelho para um flamenguista.

A XSL é uma linguagem de folha de estilos com a mesma finalidade da CSS, somente mais moderna, oferecendo mais possibilidades. Ela é composta por duas partes: XSL Transformation – XSLT responsável por transformar o documento original e a XSL – Formating Objects – XSLFO responsável pela aplicação da formatação ao arquivo transformado (RAMALHO, 2002). Observe a (FIG. 31):

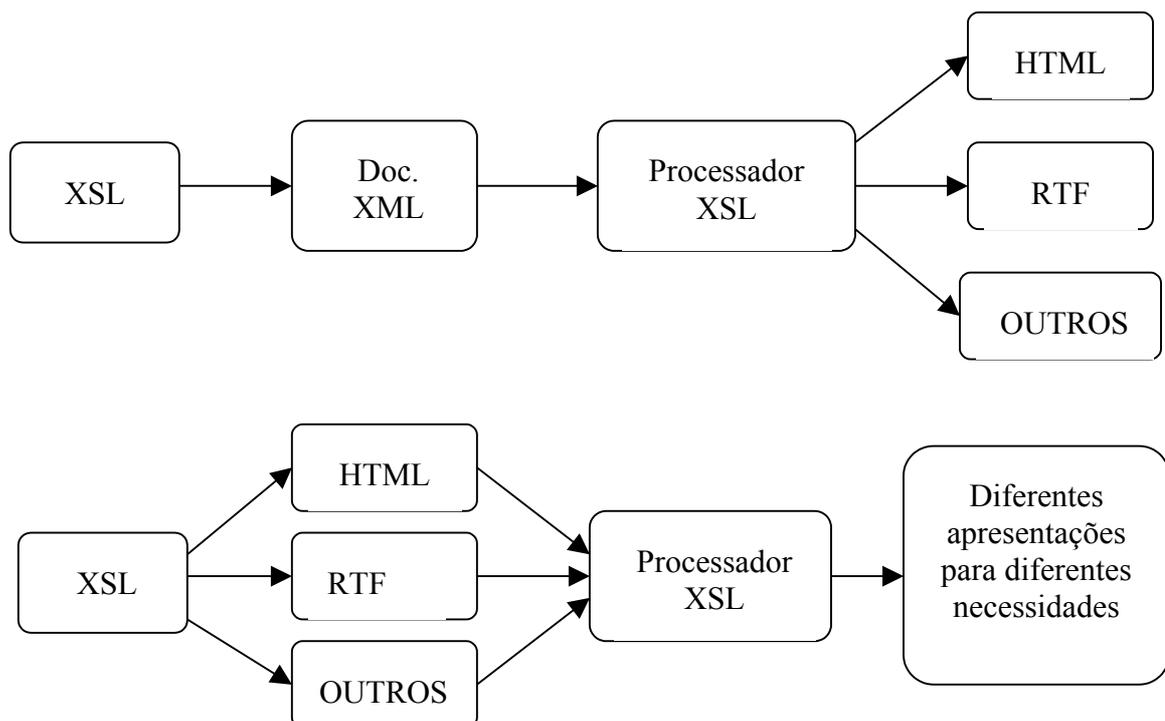


FIGURA 31: Vários estilos de apresentação à partir de um mesmo documento XML

A XSL é usada para converter um documento XML em diferentes formatos, como HTML, Rich Text Format - RTF ou qualquer outro estilo de formato existente. Para isto é necessário adicionar no início do documento XML a sintaxe exigida pela folha de estilos, assim como na DTD (FURGERI, 1999). Veja um exemplo de sua sintaxe abaixo:

```
<?xml-stylesheet href="relatorio.xsl" type="text/xsl"?>.
```

4.5.4 XLink, XPointer e XPath

XML Linking Language – Xlink é uma linguagem que permite inserir elementos em documentos XML para criar e descrever ligações entre recursos. Utiliza a sintaxe XML para criar estruturas que podem descrever hyperlinks unidirecional simples como em HTML como também descrever e criar links mais sofisticados (W3C, 2000-2003).

Para Martínez González (2000, p. 137, tradução nossa),

Os documentos podem está relacionados de diversos modos: porque sejam do mesmo autor, tratem do mesmo tema, tenham títulos similares, ou qualquer outra coisa que o usuário considere suficiente para estabelecer uma relação entre os documentos. Os links, ou ligações, permitem expressar as relações entre os dados e ampliar assim a informação do leitor com outros dados relacionados com o documento ao qual condiz no momento.

Os links foram um dos responsáveis pelo sucesso da HTML, permitindo o usuário navegar de uma página ou de um documento a outro, facilitando o seu trabalho. No entanto, os links da HTML tem algumas limitações:

[...] para retroceder ao documento de onde iniciamos a navegação é necessário recorrer ao menu da aplicação que utilizamos (no caso o navegador da Web), ou buscar no histórico da trajetória percorrida e selecionar o ponto original. Dependemos do navegador já que os links

HTML são unidirecionais, ou seja, podemos navegar desde o documento de origem até o destino URL que se deseja, mas é impossível navegar no sentido inverso (MARTÍNEZ GONZÁLEZ, 2000, p. 137, tradução nossa),

Outra limitação dos links da HTML de acordo com Martínez González (2000, p. 137, tradução nossa), que a princípio parecia não ser,

[...] é que os links HTML sempre estão incluídos no interior do documento de origem. Isto não é nenhum problema quando somos os criadores do documento, já que os inserimos nos pontos que consideramos apropriados [...]. Mas em outros casos, como os das “anotações”, é possível que queiramos comentar documentos dos quais não somos autores, e nestes portanto não teremos permissão para escrever. Esta situação é ainda mais evidente quando vários usuários compartilham um documento. A todos eles se permite “ler” o documento, mas não se permite “escrever”, desta maneira se preserva a integridade do texto original, mas não nos deixa outra solução a não ser fazer as anotações fora do documento.

Para Martínez González (2000, p. 138, tradução nossa), “estas duas limitações por si só são suficientes para justificar a necessidade de links mais potentes e flexíveis do que os que a HTML oferece”. Esse links são XML Linking Language – Xlink. E as suas características mais relevantes de acordo com a autora são:

1. **Extensibilidade:** assim como qualquer dado XML, não está limitado a uma etiqueta ou conjunto de etiquetas pré-definidas. As etiquetas de um Xlink são definidas pelo autor dos links, cuja única obrigação é indicar de algum modo que esse elemento é um link.
2. **Inclusão de informação acerca da semântica e da relação:** Com Xlink é possível incluir informação adicional sobre a relação que liga os recursos, como a data que a relação começou ou terminou, e o tipo de vínculo existente etc.

3. **Links bidirecionais:** As ligações Xlink podem ser unidirecionais ou bidirecionais.
4. **Links múltiplos:** As ligações Xlink podem ter múltiplas origens e múltiplos destinos. Por exemplo, um mesmo link pode agrupar todas as anotações feitas por distintos autores a um certo texto. Assim, uma aplicação poderia apresentar o texto e ao seu lado todas as anotações disponíveis sobre o referido texto. Também é possível o contrário, por exemplo, em um texto se pode conectar várias definições de um mesmo termo provenientes de distintas fontes, que apareceriam junto a este para que o usuário pudesse compará-las.
5. **Links fora dos documentos:** Não tem porque os links estarem obrigatoriamente dentro dos documentos (links out-of-line). Isto permite por exemplo, fazer “anotações” aos documentos sobre os quais não se tem permissão para escrever.
6. **Ligar fragmentos de documentos:** Xlink é capaz de direcionar documentos inteiros (como HTML), mas também é capaz de direcionar fragmentos concretos de um documento.
7. **Manipulação como qualquer outro dado XML:** Admitido que as ligações XLink também são dados XML, é possível explorar todas as capacidades de consulta da XML. Sobre esses links pode-se realizar consultas igual a qualquer documento XML (MARTÍNEZ GONZÁLEZ, 2000, p. 139).

Os links XML – XLink se distinguem em duas classes, de acordo com Martinez González (2000, p. 140, tradução nossa):

1. **Simples:** Unem dois recursos. São similares aos links HTML.
2. **Estendidos:** Não tem equivalente em HTML. Englobam os links múltiplos que unem mais de dois recursos e out-of-line.

Existe um padrão associado ao XLink que aborda o aspecto de direcionamento de fragmentos internos dos documentos, este se chama XML Pointer Language – Xpointer, que está baseado em XML Path Language – XPath. Com o recurso da XPath uma parte de um documento se direciona com base no caminho que ele tem que seguir para se chegar a um outro documento ou recurso.

4.5.5 XML Namespaces

A linguagem XML permite ao desenvolvedor criar seus próprios elementos, com nomes e atributos, aliás, uma das grandes vantagens da XML. Porém, isso pode gerar alguns problemas quando uma aplicação XML, desenvolvida por várias pessoas ou instituições, gera nomes de elementos ou atributos duplicados (RAMALHO, 2002). Qin (2000, p. 86, tradução nossa) salienta que “para complicar ainda mais o assunto, futuros documentos XML conterão múltiplos vocabulários de marcação e isso causará problemas para reconhecimento, gerando conflitos”.

Por exemplo, um elemento <titulo> pode se referir a coisas diferentes como, o título de um livro, o título de um artigo de jornal ou o título de uma revista entre outras coisas. Para resolver esse problema usa-se um prefixo para cada elemento <titulo> a fim de identificá-lo corretamente. Por exemplo: <livro: titulo>; <artigodejornal: titulo>; <revista: titulo>. Os nomes dos elementos relacionados no interior das tags, são denominados **namespaces** (RAMALHO, 2002), “um conjunto de nomes de elementos relacionados” (CASTRO, 2001, p. 113), cuja “função principal [...] é identificar de forma única os elementos e atributos de um documento XML” (RAMALHO 2002, p. 140).

No entanto, de acordo com Qin (2000, p. 86, tradução nossa),

A solução para os problemas de namespaces de XML estará em grande parte nas mãos da comunidade bibliotecária e da ciência da informação que, durante anos de pesquisas em organização e representação da informação e do conhecimento, tem desenvolvido uma estrutura inteira de metodologias e sistemas. Um exemplo imediato é que as técnicas usadas na criação e controle de tesouros podem ser aplicadas para padronizar as nomeações ou marcações dos elementos de dados em várias aplicações XML e traçar a semântica dos mesmos nos repositórios de namespaces.

Para não ter problemas com a duplicação de namespaces, no processo de troca de informações, é preciso criar normas com o objetivo de torná-los universalmente únicos e assim evitar conflitos, pois, cada vez mais, “surgem aplicações XML, e a demanda para controle e gerenciamento de namespaces também aumentará” (QIN, 2000, p. 86, tradução nossa), visto que uma das exigências das organizações de hoje é ter sistemas de informação efetivos que possam responder rapidamente as necessidades de informação para gerar tomadas de decisões mais seguras e eficientes.

Veja na figura 32 um trecho do namespace proposto por Lam (2002, p. 91), Chefe de Sistemas da Biblioteca da Universidade de Ciência e Tecnologia de Hong Kong:

```
<?xml version= "1.0" encoding= "UTF-8"?>
<marc>
  <fd name= "245" ind1= "1" ind2= "4" ... other attributes... >
    <sf name= "a"> The ABCs of XML : </sf>
    <sf name= "b"> the librarian's guide to the extensible markup language / </sf>
    <sf name= "c"> by Norman Desmarais </sf>
  </fd>
  ... more fd elements follow ....
</marc>
```

FIGURA 32: Proposed namespace

FONTE: LAM, 2002, p. 91

No exemplo da (FIG. 32),

`<marc>` e `</marc>` marcam o início e o término de um registro MARC. `<fd>` e `</fd>` marcam o início e o término das etiquetas MARC. `<sf>` e `</sf>` marcam o início e o término de um subcampo. Os elementos **fd** e **sf** contêm vários atributos. O atributo **name** define o nome da etiqueta MARC ou o código de subcampo. **ind1** e **ind2** são os dois indicadores da etiqueta MARC (LAM, 2002, p. 91, tradução nossa).

4.6 Uso da XML

Devido as limitações dos Browsers existentes atualmente para a visualização de páginas XML, o uso real desta linguagem ainda não se concretizou de fato. Por exemplo, os navegadores mais antigos simplesmente não reconhecem a linguagem, tratando os arquivos em XML como estranhos blocos de texto (CASTRO, 2001).

Segundo Castro (2001, p.15),

O maior empecilho para difundir páginas em XML é que nenhum navegador é compatível com Xlink ou Xpointer. E isso significa que nenhum navegador pode visualizar links ou imagens em uma página XML. Até que isso seja resolvido, ninguém vai oferecer diretamente páginas em XML.

Para contornar esse problema temporariamente, usa-se a XML para gerenciar e organizar as informações e depois aplica-se a XSLT para converter esses documentos XML em HTML para a possível visualização em um navegador (CASTRO, 2001).

“O World Wide Web Consortium (W3C) recomenda utilizar XHTML – um sistema usado para escrever tags HTML de acordo com as estritas regras da XML” (CASTRO, 2001, p. 15).

Para Castro (2001) a XHTML é problemática pelo seguinte fato: perde-se a natureza amigável da HTML mas não ganha o poder de identificar informações da XML.

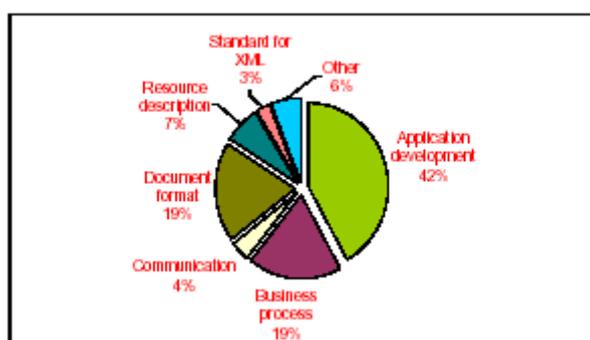
A XML está sendo considerada por muitos como a sucessora da HTML na Internet. Embora já tenham feito implementações nos servidores web e navegadores que suportam XML ainda há muito trabalho a ser feito até que estas aplicações sejam as mais usuais. Agora mesmo grande parte do esforço se concentra em conseguir que os dados XML sejam acessíveis pelo máximo de aplicações, para assim difundir sua utilização. Este é o trabalho que realizam as ferramentas que se podem acoplar aos servidores web, que permitem transformar os dados XML em HTML. Essas ferramentas são processadores de XML e XSL (processadores XSLT) (MARTÍNEZ GONZALEZ, 2000, p. 136, tradução nossa).

No que diz respeito aos navegadores, estes já estão manipulando XML. Alguns dos navegadores que já suportam a XML são, Internet Explorer 6 e o Netscape Communicator 6 entre outros. O Internet Explorer 5 também suporta a XML mas com um conjunto limitado: folhas de estilos CSS e uma versão reduzida de XSL (uma versão obsoleta).

Contudo, são vários os usos da XML. Segundo Furgeri (2001) são diversas as possibilidades de utilização da XML:

- Para melhorar a busca de informações;
- Para automação da Web;
- Para troca de dados na Web;
- Ser usada em conjunto com documentos HTML;
- Para desenvolvimento de catálogos;
- Para automação da cadeia de suprimentos;
- Para publicações científicas;
- Aplicações comerciais entre muitas outras possibilidades de uso.

De acordo com Qin (2000, p. 86, tradução nossa) em sua pesquisa realizada em outubro de 1999, sobre as iniciativas e aplicações emergentes da XML constatou-se que elas variam de processos empresariais ou de negócios às disciplinas científicas. As diversas áreas de aplicações XML, segundo o autor, podem ser visualizadas na figura 33:



Desenvolvimento de aplicações - 42%
 Processos empresariais ou de negócios - 19%
 Comunicação - 4%
 Formatos de documentos - 19%
 Descrição de recursos - 7%
 Padrões para XML - 3%
 Outras - 6%

FIGURA 33: Áreas de aplicações XML
FONTE: QIN, 2000, p. 86

Nesta mesma pesquisa do autor Qin (2000, p 86, tradução nossa) foi constatado que as empresas e as associações de indústrias são as fomentadoras mais ativas em iniciativas de aplicações XML. Observe a figura 34:



Organizações envolvidas em aplicações XML

Empresas/Negócios - 42%
 Governo - 3%
Associações de indústrias - 30%
 Institutos de pesquisas - 4%
 Sociedade escolar - 4%
 Universidades - 4%
 Outras - 13%

FIGURA 34: Categorias de organizações envolvidas em aplicações XML.
FONTE: QIN, 2000, p. 86

É interessante notar que esses dados estatísticos foram obtidos em 1999, logo após a liberação oficial da XML em 1998 pelo W3C. Não encontramos pesquisa com dados mais

recentes, mas calcula-se que os números estatísticos do uso da XML atualmente pelas instituições estejam bem altos, devido a sua enorme aceitação.

Segundo Lam (2002, p. 90, tradução nossa),

Desde a sua liberação oficial em 1998 pelo W3C, a XML foi rapidamente adotada e implementada por muitas indústrias. Nos últimos anos, muitas iniciativas e projetos que fazem uso da XML para metadados de biblioteca emergiram (dados bibliográficos, controle de autoridades, cabeçalhos de assunto e outros).

Como podemos perceber a XML hoje é não somente uma realidade mas uma necessidade. É a promessa de tornar a Internet mais “inteligente”, transformando o modo de armazenamento, recuperação e intercâmbio de informações e, principalmente, a economia da informação na forma de redução de custos operacionais (FURGERI, 2001) para as instituições, e desta forma causando grandes mudanças em diversas áreas, tornado-se necessário examinar o impacto que esta vem causando, inclusive no mundo bibliotecário.

Para Ortiz-Repiso Jiménez (1999, p. 222, tradução nossa) “tanto [...] [XML] como MARC são linguagens de marcação descritiva que criam textos estruturados”. Mas a XML permite flexibilidades que não podem ser encontradas no MARC.

Estaremos abordando os dois esquemas de marcação no próximo capítulo Estruturas de Marcação MARC21 e XML.

CAPÍTULO 5

ESTRUTURAS DE MARCAÇÃO MARC21 e XML

Quando as bibliotecas começaram a usar o formato MARC para seus catálogos nos anos 1960, elas converteram seus registros impressos para a forma eletrônica para armazenamento e recuperação. Os materiais representados por estes registros são físicos e estáticos. No espaço da Web, não há muito físico, nem estático, o material é virtual e a informação é dinâmica.

Jian Qin

O MARC é considerado a primeira estrutura e padrão criados para a construção e representação de metadados produzidos na atualidade, ou seja, dados bibliográficos e catalográficos, com o objetivo de possibilitar a troca desses dados entre bibliotecas utilizando-se de computadores (GRÁCIO, 2002; SOUZA; CATARINO; SANTOS, 1997). Desenvolvido na década de 1960 originalmente para descrever livros impressos, com o passar dos anos foi sofrendo adaptações para proporcionar descrição a uma variedade de tipos de documentos, assim como o acesso e localização da informação nos recursos da Web.

O MARC foi a ferramenta que possibilitou a concretização da Catalogação Cooperativa em ambiente digital e em rede, e serviu de base para outros padrões para metadados criados na atualidade (Government Information Locator Service – GILS; Dublin Core – DC; Content Standard for Digital Geospatial Metadata – CSDGM, entre outros). “Hoje o formato continua a ser usado como formato de intercâmbio e comunicação de dados e todos os grandes sistemas de catalogação suportam a importação e exportação de registros no formato MARC” (FIANDER, 2001, p. 18, tradução nossa).

É uma família de formatos, Bibliográfico, Autoridade, Cabeçalho de assunto, Classificação e Informação de comunidade. Dá suporte há mais de trinta anos à comunidade bibliotecária e a indústria de software especializado.

A sua rica contribuição à padronização e à comunidade por redes é inquestionável. Porém, devido a sua estrutura rígida e complexa, projetada para atender as necessidades de uma época em que os catálogos impressos ou em microfimes eram a norma, tornou-se limitada para atender as diferentes necessidades informacionais de hoje (ROWLEY, 1994).

De acordo com Ortiz-Repiso Jiménez (1999, p. 222, tradução nossa) “o formato MARC que consolidou as bases de cooperação e comunicação de informação bibliográfica no mundo bibliotecário tem, porém, uma série de características negativas” e limitações para atender as atuais demandas informacionais.

Para Ortiz-Repiso Jiménez (1999, p. 222, tradução nossa) essas características poderiam ser sintetizadas da seguinte forma :

- Está estritamente controlado, qualquer mudança ou acréscimo no formato demora anos para ser realizado. Por exemplo, o sinal gráfico “til” (~) que tanto se utiliza nas URLs das páginas da Web, demorou mais de dois anos para incorporar-se no jogo de caracteres do [...] [MARC]. O projeto InterCat da OCLC não podia operar com URLs que tivesse esse caractere;
- É trabalhoso, lento e custoso de realizá-lo e devido a sua complexidade deve ser manipulado por profissionais qualificados.
- Embora seja composto por campos de longitude variável, está limitado a uma longitude máxima de 100.000 caracteres que, é mais que suficiente para registros bibliográficos tradicionais de bibliotecas, só não é para outros tipos de registros de outros âmbitos (arquivos, por exemplo).
- Se adapta muito mal a informação estruturada hierarquicamente já que se baseia numa estrutura plana. Foi projetado para descrever e permitir a

informação de um registro bibliográfico e não para estabelecer relações entre registros.

Segundo a autora Ortiz-Repiso Jiménez,

O formato MARC tem se mantido fundamentalmente como uma versão eletrônica do catálogo em fichas com as limitações próprias deste modelo. A principal função deste formato quando desenhado pela primeira vez nos anos 1960, foi permitir a distribuição eletrônica de registros bibliográficos para a produção de fichas em papel. Ao formato MARC falta, em definitivo, habilidade suficiente para tratar a informação estruturada hierarquicamente e proporcionar o acesso a coleções completas [...] (ORTIZ-REPISO JIMÉNEZ, 1999, p. 223, tradução nossa).

Para Tennant (2002a, p.1, tradução nossa) “no MARC, os campos não são explicitamente rotulados, mas codificados com um esquema numérico que não pode ser lido por alguém desfamiliarizado com a sintaxe complexa”.

Tennant coloca em pauta a questão da marginalização do formato, visto que na opinião do autor “o MARC tem sido sempre um padrão secreto, nenhuma outra profissão usa o MARC ou qualquer coisa parecida com ele” (TENNANT, 2002a, p.2, tradução nossa). E os softwares para o formato são limitados, deixando muitas vezes as instituições e os seus profissionais sem escolha, uma vez que estes estão restritos a um nicho de mercado de vendedores de sistemas para bibliotecas (TENNANT, 2002a).

Segundo Johnson (2001, p. 82, tradução nossa),

A equipe de profissionais da Lane Library [Biblioteca Médica Lane da Universidade de Stanford] examinou o formato bibliográfico MARC21 e descobriu que, em seu ponto de vista, os formatos MARC prejudicam uma integração efetiva de dados nas páginas da Web. Especificamente pelo fato de que o MARC emprega uma estrutura e linguagem de marcação diferente da XML e da HTML, sendo ambas os padrões mais próximos identificados

com o desenvolvimento da Web. O uso contínuo do MARC garantirá que algumas informações bibliotecárias serão úteis na Web e outras não. Dick Miller chamou a isso de “Segregação Bibliográfica”.

Dos problemas específicos apontados pelo pessoal da Lane Library em relação ao MARC21 segundo Johnson (2001, p. 83, tradução nossa), incluem:

- A informação é registrada de diversas maneiras, as vezes inconsistentes, e em diversos lugares. Isso pode ser observado na forma em que a informação de data é registrada. Algumas datas, por exemplo, são encontradas nos registros MARC21 nos campos 005, 008/00-05, 008/07-10, 008/11-14, e 260 subcampo “c”, em quatro formatos distintamente diferentes.
- Há complexidades inerentes no MARC que são controladas incompativelmente pelos diferentes formatos. Por exemplo, é embutida informação do tipo campo fixo nos campos 4XX, 5XX, e 7XX subcampo “w” em registros de autoridade, e nos campos 76X, 78X subcampo “w” nos registros bibliográficos.
- Os registros MARC, incluem informações registradas em forma codificada que, na visão do Sr. Miller⁴, não pode ser justificada no ambiente computacional de hoje com muitas mídias de armazenamento mais baratas. Exemplos de informações que em sua opinião são registradas desnecessariamente em forma codificada, incluem: publicações governamentais (008), idioma (008, 041), país (008, 043) e outros.
- Finalmente, o formato MARC é visto como “plano, com suporte limitado para hierarquia” e, em sua opinião não suporta bem os links da web.

Em relação a hierarquia das informações, Tennant, (2002b, p. 1, tradução nossa) explica claramente que:

⁴ Chefe dos Departamentos de Sistemas Técnicos e Sistemas Bibliotecários da Stanford University’s Lane Medical Library - Biblioteca Médica Lane da Universidade de Stanford.

Algumas informações são hierárquicas por natureza, enquanto outras não. A maioria das informações bibliográficas não tem relações hierárquicas, ou seja, um livro pode ter um ou mais autores, um título, alguns cabeçalhos de assunto, etc. Já nenhuma dessas informações podem ser logicamente abrigadas com outras informações. Os conteúdos de alguns livros, por outro lado, podem ser completamente hierárquicos por natureza, uma seção pode ter um ou mais capítulos nela contidos.

Segundo Fiander (2001, p. 26, tradução nossa) “a rigidez e as irregularidades internas do MARC tem começado a causar problemas para catalogadores e usuários, e está começando a ficar para trás em pesquisas atuais sobre padrões de descrições bibliográficas”.

Mediante as características negativas e as limitações do MARC citadas anteriormente, de acordo com Ortiz-Repiso Jiménez (1999, p. 222, tradução nossa), “[...] [a XML] proporciona uma qualidade de trabalho promissor por diversas razões”:

- Pode tratar informações hierarquicamente inter-relacionadas em tantos níveis quantos forem necessários.
- Não tem limitações para os tamanhos dos documentos;
- É um padrão internacional adotado por um número crescente de instituições governamentais, de pesquisa e da indústria.

Segundo Ortiz-Repiso Jiménez,

Tanto [...] a [XML] como o MARC são linguagens de marcas descritivas que criam textos estruturados. No entanto, a [...] [XML] permite uma flexibilidade máxima no uso do texto. O usuário pode controlar os formatos de indexação, apresentação e impressão. Sua estrutura facilita a criação de bases de dados mais sofisticadas que as baseadas em MARC e permite a indexação e a recuperação do documento, como documento e, também, os componentes do documento mediante busca booleana, adjacência, proximidade, ranking de relevância e outros. A estrutura do texto em [...]

[XML] suporta também navegação em linha avançada (ORTIZ-REPISO JIMÉNEZ, 1999, p. 222, tradução nossa).

Uma das necessidades informacionais de hoje é a flexibilidade da informação, e “em definitivo a flexibilidade do texto estruturado em [...] [XML] é inquestionavelmente superior a hierarquia plana do texto estruturado em MARC”. (ORTIZ-REPISO JIMÉNEZ, 1999, p. 222, tradução nossa).

De acordo com a autora,

A [...] [XML] pode ser usada para criar uma grande variedade de documentos e usá-los também para proporcionar acesso e controle a uma grande quantidade de formatos de informação em linha. Não é uma linguagem de marca proprietária, como o formato MARC e não depende tão pouco de um software proprietário ou de uma oferta limitada de software. Por natureza, poderia ser possível usar a [...] [XML] como um padrão geral para as ferramentas informativas que se usam para catalogar (Regras, formato MARC, CDU e outros), para criar os registros catalográficos e para criar os textos eletrônicos do catálogo. E também poderíamos utilizar o texto para proporcionar acesso e controle aos objetos digitais que não são texto (ORTIZ-REPISO JIMÉNEZ, 1999, p. 223, tradução nossa).

O Sr. Miller citado por Johnson (2001, p. 84, tradução nossa),

- Caracteriza a XML como o “formato de dados universal para a Web”, uma “SGML simplificada, usando a sintaxe HTML, com facilidades da Web acrescentadas”.
- Emprega um conjunto único de caracteres, Unicode para representar todos os dados.

O Sr. Miller, de acordo com Johnson (2001, p. 84, tradução nossa) também observou que, “um aspecto significante da XML é a sua separação do conteúdo, apresentação e ligação, de forma que cada um possa ser controlado otimamente.

Segundo Bax (2002) esta é a filosofia da XML (a separação da apresentação, do conteúdo e da ligação). O formatador na FIG. 35 está representando a XSL, que possibilita diferentes apresentações de um mesmo documento XML, conforme explicado no capítulo anterior. Veja a figura :

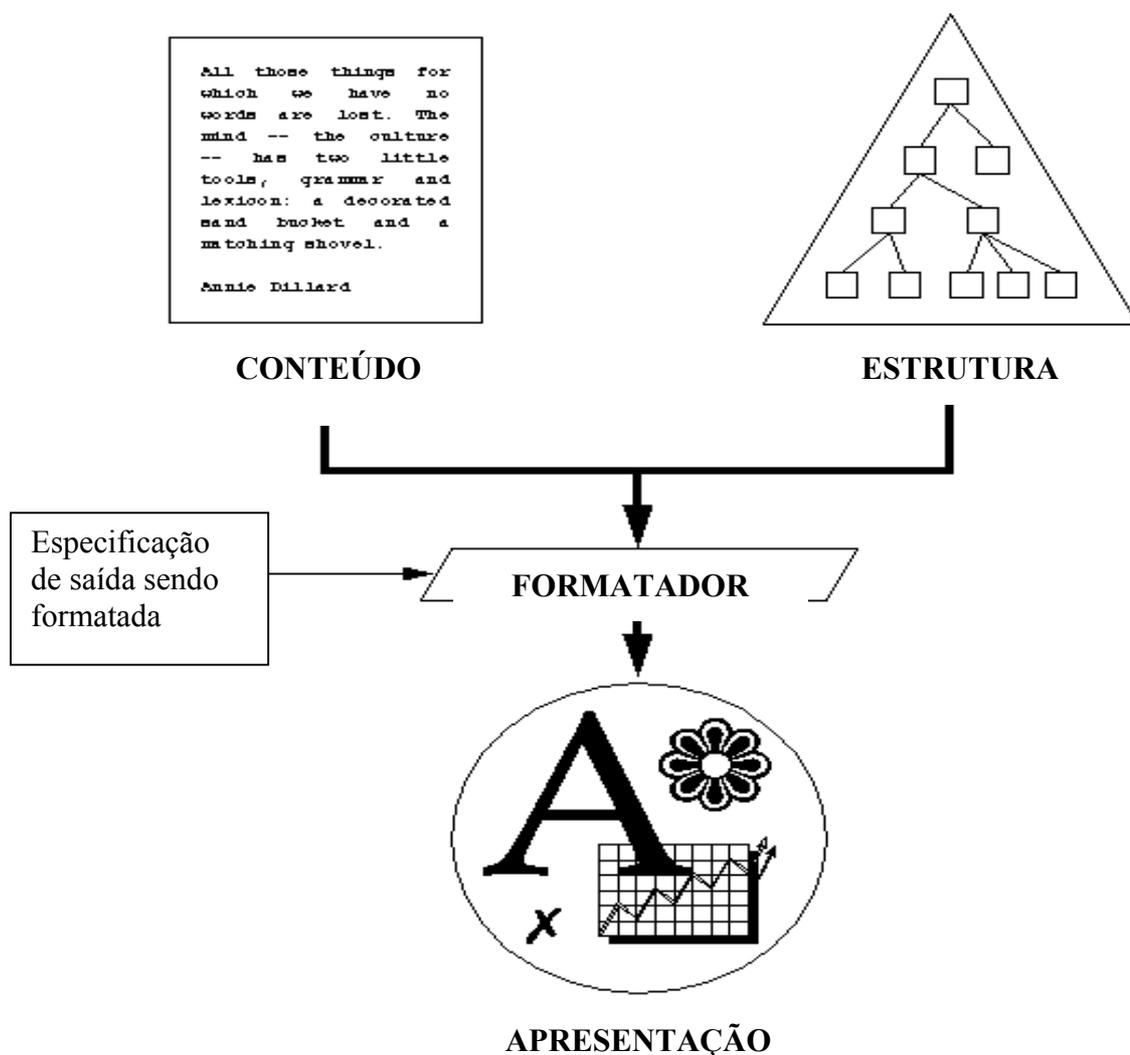


FIGURA 35- Filosofia da XML.
 FONTE- Bax, 2002, slide 18, tradução nossa.

Para Bax (2002, slide 13) a XML é:

- Padrão para se criar padrões
 - Formato universal para troca de dados e documentos
 - Espécie de “língua franca” da era da informação
 - Poder, precisão, controle e flexibilidade, atraindo a atenção dos bibliotecários
- Padrões abertos (*open source movement*)
 - Não proprietário – qualquer um pode usar
- Extensibilidade (relativamente fácil de expandir)
- Separação entre conteúdo e apresentação
 - Mantém a flexibilidade da informação
- Técnicas enlaces (links) mais sofisticadas
 - Linkar uma parte específica do documento
- Neutralidade de plataforma e software / Interoperabilidade
- Unicode: longevidade, internacionalização
 - Conjunto de caracteres únicos para todas as línguas humanas
- Interface com banco de dados

De acordo com Brun (2002, slide 4, tradução nossa) a XML:

- Oferece um sistema de codificação mais simples e fácil de processar que o utilizado atualmente para registros MARC (2709);
- Apresenta-se como um formato “universal”, utilizado em qualquer tipo de aplicação, frente ao sistema de codificação atual do MARC;
- Elimina a complexidade desnecessária para os desenvolvedores;
- Pretende evitar a “segregação” do MARC frente a outros formatos de metadados para a Internet utilizados nos projetos de “biblioteca eletrônica”.

Segundo Brun (2002, slide 3, tradução nossa) “o formato XML tem se apresentado como uma alternativa para representar e transferir registros MARC através da Internet e facilitar seu tratamento automatizado”.

A união da estrutura de representações bibliográficas e catalográficas MARC e a linguagem XML será extremamente valiosa ao considerarmos as diferenças entre as estruturas de marcação MARC e XML conforme apresentou Bax (2002):

MARC	XML
<p>Estrutura Plana O MARC é bidimensional (campo/subcampo).</p>	<p>Estrutura Hierárquica A XML permite infinitas possibilidades de aninhamento.</p>
<p>Formato proprietário Restrições impostas à sua utilização</p>	<p>Formato aberto Sem restrições. Qualquer pessoa ou instituição pode utilizá-la. Autonomia para a criação de DTDs, para necessidades específicas e marcações personalizadas.</p>
<p>Dependência de plataforma O MARC de uma certa maneira sempre foi proprietário, dependente de plataforma específica.</p>	<p>Independência de plataforma A XML é independente de plataforma, o software capaz de manipulá-la é o próprio navegador.</p>
<p>Esquema de marcação De difícil compreensão para pessoas desfamiliarizada. Esquema codificado.</p>	<p>Esquema de marcação Compreensível por humanos e máquinas. Esquema semântico.</p>

MARC	XML
<p>Acesso Restrito Quase que 100% restrito à comunidade bibliotecária.</p>	<p>Acesso Aberto Enquanto que todos os outros profissionais da informação estão se aprofundando no uso da XML, além desta ser acessível a qualquer pessoa, qualquer área e a qualquer profissional.</p>
<p>Registros tradicionais Limitado aos registros tradicionais.</p>	<p>Registros novos A XML pode representar registros que não existiam: Textos longos (texto completo, biografia, resumos –gráficos). Em XML encontramos os dados primários (o próprio documento) juntamente com dados secundários – (meta)dados.</p>
<p>Sintaxe Em geral o MARC é uma sintaxe + um compromisso com uma implementação (um modelo de dados, limites de tamanhos específicos 99.999) Tendência ao desenvolvimento proprietário.</p>	<p>Sintaxe A XML é apenas uma sintaxe - separação mais clara entre a semântica da representação do registro e a semântica da aplicação.</p>

FIGURA 36- Esquema MARC e XML

Uma das grandes vantagens da XML é a de ser uma estrutura de marcação aberta, ou seja, não impõe restrições à sua utilização nem a criação de ferramentas que a manipulem. Segundo Shapiro e Varian (1999) com um padrão aberto as Instituições não se preocuparão

com o aprisionamento por soluções proprietárias. Veja na Figura 37 o ciclo do aprisionamento segundo os autores.

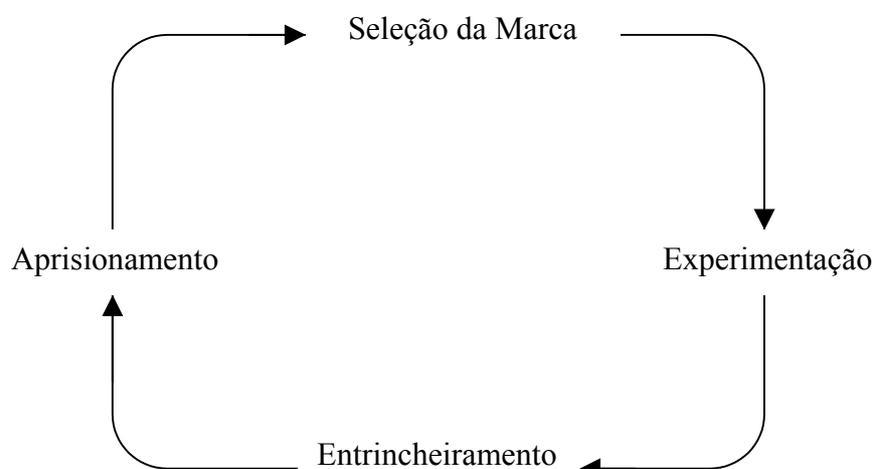


FIGURA 37 – O ciclo do aprisionamento
FONTE: Shapiro e Varian (1999, p. 158)

De acordo com Shapiro e Varian (1999, p. 158) “o lugar mais fácil para lançar-se no ciclo do aprisionamento é no ponto de seleção da marca”, ou seja, a escolha de um novo produto, por exemplo, um novo programa de software ou uma nova base de dados. De acordo com os autores “não nascemos “aprisionados”, só ficamos em virtude das escolhas que fazemos” (SHAPIRO ; VARIAN, 1999, p.158).

Depois da seleção da marca “segue-se a fase de experimentação durante a qual o cliente usa de maneira ativa a nova marca e usufrui da vantagem de todos os incentivos que teve para dar uma oportunidade a essa marca” (SHAPIRO; VARIAN, 1999, p.158). Por exemplo, as ofertas grátis de produtos como softwares livres para uso 30, 60 ou 90 dias.

Os clientes que fazem mais do que experimentar entram para a fase de entrincheiramento. Isto é, quando o consumidor [ou instituição] realmente acostuma-se à nova marca [ou novo produto], desenvolve uma preferência por ela em detrimento das outras e talvez fique retido a essa marca ao fazer os investimentos complementares. Normalmente o fornecedor tenta prolongar essa fase e adiar a consideração ativa de outras marcas [ou

produtos], desejando que os custos de troca do cliente subam. A fase de entrincheiramento culmina com o aprisionamento quando os custos de troca tornam-se altos demais. (SHAPIRO; VARIAN, 1999, p.159).

“A essência do aprisionamento é que suas escolhas no futuro serão limitadas pelos seus investimentos de hoje” (SHAPIRO; VARIAN, 1999, p. 161).

Segundo os autores existem 7 (sete) padrões econômicos primários que levam ao aprisionamento: “Compromissos contratuais, equipamentos duráveis e atendimento pós-venda, treinamento específico de marca, informação e banco de dados, fornecedores especializados, custos de busca e programas de lealdade” (SHAPIRO ; VARIAN, 1999, p. 161).

Agora, mesmo os padrões abertos tem algumas desvantagens, por exemplo, segundo Shapiro e Varian (1999, p. 294), eles “são propensos a “estilhaçamento” ou “fragmentação” e podem também serem “seqüestrados”. De acordo com Shapiro e varian (1999, p. 294) “o estilhaçamento de um padrão refere-se ao surgimento de versões múltiplas e incompatíveis de uma tecnologia padronizada”.

O exemplo clássico dos riscos de administrar padrões abertos e dos riscos do estilhaçamento, é a história do sistema operacional Unix. O UNIX foi desenvolvido originalmente nos Laboratórios Bell como uma ferramenta de pesquisa. A AT&T cedeu o código fonte a pesquisadores acadêmicos por muitos anos, e ele tornou-se um padrão na comunidade de pesquisas.

Quando o mercado de minicomputadores decolou na década de 1970, o UNIX foi modificado e vendido por diferentes empresas; a disseminação acelerada das estações de trabalho na década de 1980 conduziu a mais versões do UNIX, e nenhum padrão setorial foi estabelecido (SHAPIRO; VARIAN, 1999, p.294).

De acordo com Shapiro e Varian (1999, p. 296),

Os padrões abertos também podem ser “seqüestrados” por empresas que buscam estendê-los em direções proprietárias e, portanto, ganhar controle sobre a base instalada com o tempo. A Microsoft tem sido acusada de tentar estender tanto a Java como a HTML em direções proprietárias.

São por estas razões que os padrões abertos necessitam de responsáveis que possam administrá-los, e que sejam encarregados de estabelecer a direção na qual o padrão evoluirá e mantê-los livres da estagnação, fazendo investimentos em recursos para o seu desenvolvimento e manutenção.

Segundo Shapiro e Varian (1999, p. 298) “Sem um responsável, os padrões abertos podem estagnar-se ou estilhaçar-se em pedaços incompatíveis”.

Quanto a XML o órgão responsável pelo seu desenvolvimento e manutenção é o W3C, como já foi explicado no capítulo anterior.

Alguns autores como Amanda Xu (1997) citado por Ortiz-Repiso Jiménez (1999, p. 224, tradução nossa),

considera que o formato MARC continua sendo o mais apropriado por sua estrutura e sobre tudo porque seu uso está muito estendido. Propõem criar programas de conversão entre os distintos formatos de metadados com o objetivo de poder incluir os registros nos catálogos bibliotecários.

Já a autora Virgínia Ortiz-Repiso Jiménez (1999, p. 224, tradução nossa) diz:

Eu não considero que o MARC seja a melhor forma de descrever metadados, o que seria interessante converter os metadados em outros formatos a convertê-los à MARC. Penso que podemos nos encontrar na mesma situação que quando se instalava um sistema integrado de gestão bibliotecária e se seguia mantendo o catálogo manual. Chegará um momento em que somente um dos dois sistemas prevalecerá.

Segundo Ortiz-Repiso Jiménez (1999, p. 224, tradução nossa),

O catálogo do futuro deverá ser flexível e multitarefa e o formato MARC não pode proporcionar essa flexibilidade. Isto não quer dizer que o formato MARC já não sirva. Há bilhões e bilhões de registros MARC e o custo e tempo de convertê-los para [...] [XML] seria esmagador. Mas eu acredito que o formato MARC não será por muito mais tempo o único formato de codificação de dados bibliográficos nos sistemas bibliotecários. Pois o formato MARC não é suficientemente flexível para permitir que as bibliotecas aproveitem a tecnologia de novos desenvolvimentos de recuperação e acesso a informação.

Porém, Khurshid (2002, p. 370, tradução nossa) aponta que,

Apesar da avaliação de alguns bibliotecários [que trabalham com informação digital] de que o formato MARC provê uma pobre adaptação para catalogar recursos eletrônicos, ele ainda é usado nos projetos digitais por duas razões:

- (1) Ele carrega muita informação padronizada, é de fácil processamento, as seqüências de bytes são claramente designadas, e
- (2) Outros padrões para codificar informação digital em forma legível por máquina como Dublin Core, Extensible Mark-up Language (XML) e Resource Description Framework (RDF) ainda estão em desenvolvimento.

Para Tennant (2002b, p. 3, tradução nossa) existem algumas saídas estratégicas para o MARC:

Um método é o **enterramento**. Minha opção menos favorita é deixar de criar novos registros MARC e começar a construir um sistema paralelo baseado em um novo padrão. Por meio das ferramentas de busca de bancos de dados relacionais que se tornaram uma realidade, pode ser possível criar

um portal que costure os dois sistemas para o usuário enquanto os profissionais da informação trabalham para migrar os registros antigos com o passar do tempo.

Outra opção é a **encapsulação**. A Biblioteca do Congresso já está trabalhando para definir um padrão para informação bibliográfica em XML que quebre algumas marcas do MARC. Este padrão em desenvolvimento, o **Metadata Object Description Schema (MODS)**, é suficiente para capturar a maioria do que está agora em registros MARC, embora não seja possível a "viagem de ida-e-volta" (ir do MARC para MODS e retornar ao MARC novamente) com registros MARC usando MODS.

Segundo o autor, para um projeto na Biblioteca Digital da Califórnia, eles extraíram registros MARC de seus catálogos, os traduziram em MODS, e embutiram os registros em uma envoltura XML através do **Metadata Encoding e Transfere Sintaxe (METS)**. Demonstraram também como poderiam pegar metadados do editor e embutí-los numa seção própria do registro METS. Eles puderam então indexar e exibir seletivamente campos de qualquer fonte de metadados para o mesmo item.

E uma outra saída estratégica para o MARC segundo Tennant (2002b, p. 4, tradução nossa) “seria a **migração**. Uma transição completa para um novo jogo de padrões bibliográficos requereria migrar a base de registros existentes e tudo que nós [profissionais da informação] inventamos”.

Segundo Tennant (2002b, p. 4, tradução nossa) em relação a migração ele sugere que:

Antes de contestar tal empreendimento, é preciso considerar que há alguns anos atrás não havia nenhum registro, nada na forma digital. E migrar informação digital de um formato para outro será mais fácil, desde que algumas partes (mas provavelmente não tudo) possam ser automatizadas. O objetivo é fazer esta migração uma vez e para um padrão que seja flexível o

bastante para permitir extensões futuras e mudanças sem requerer outra migração.

Porém, de acordo com Khurshid (2002, p. 376, tradução nossa) :

Sugerir que o MARC deve ser substituído por esquemas de metadados não é aceitável por uma vasta maioria da comunidade bibliotecária. Ao invés disso eles sugerem que deve ser permitido a coexistência do MARC e dos esquemas para metadados. A LC, OCLC e outras organizações já aceitaram a importância disto e estão atualmente engajadas no desenvolvimento de mapas e caminhos cruzados, de forma que dados possam ser intercambiados.

No entanto, várias bibliotecas já iniciaram a migração e o uso do formato MARC em ambiente XML como, a Medlane⁵ com o formato **XMLMARC**⁶ e a Library of Congress com o **MARCXML**⁷, entre outras.

O **MARCXML** não será abordado aqui neste trabalho, ficando portanto a recomendação para uma futura pesquisa já que é um tema rico em informações e extremamente necessário para a área da Biblioteconomia e da Ciência da Informação.

⁵ Stanford University's Lane Medical Library

⁶ <http://xmlmarc.stanford.edu>

⁷ <http://www.loc.gov/standards/marcxml>

CAPÍTULO 6

CONSIDERAÇÕES FINAIS

É necessário levar para todos os espaços possíveis, a idéia da mudança.
Mudança de paradigmas.

Gramigna

A representação e organização da informação, na era digital, assim como os negócios e a indústria, tiveram mudanças dramáticas em quase todas as fases de seus processos. As mudanças incluem não somente os métodos e a tecnologia usada na criação de formas de representação de documentos e informações, mas também nos padrões, que são essenciais para o sucesso e a efetividade das ferramentas de busca e recuperação da informação.

Diante deste novo cenário, fundamentalmente enraizado em ambientes e infraestruturas para a informação digital, o coração deste novo espaço informacional está na existência de padrões que possibilitam a diferentes tipos de dados serem intercambiados e entendidos por plataformas e sistemas heterogêneos. (QIN, 2000). Podemos resumir tudo isso em uma única palavra: interoperabilidade, ou seja, compatibilidade entre diferentes sistemas e/ou plataformas.

O uso da XML nos padrões de representação documentária nos possibilita, além da interoperabilidade, muitas outras facilidades e flexibilidades tanto na representação quanto na organização e na recuperação da informação. Podemos perceber que as inovações tecnológicas têm como objetivo facilitar, inovar e dar novas oportunidades a todos os seguimentos da sociedade, como um benefício.

Segundo Gazan, (2000, p. 14, tradução nossa) ,

O formato MARC [...] tem permitido às instituições criarem e intercambiarem dados bibliográficos efetivamente por décadas. A XML é uma extensão deste ideal. Ela possui um jogo de padrões abertos para expressar dados como um simples arquivo de texto de campos e valores. E ainda realiza aparentemente o impossível: Deixa grupos e indivíduos livres para definirem as suas próprias linguagens de marcação de dados especializados, enquanto ainda provêem os meios para a interoperabilidade entre linguagens.

A XML, em relação à Biblioteconomia, poderia ser considerada como uma inovação tecnológica para o MARC21, vindo a se posicionar num futuro próximo como uma nova tecnologia, ou como o redesenho de uma série de tecnologias já existentes como SGML, HTML, MARC e outras, dando como resultado o desaparecimento do MARC21 ou a coexistência deste com outros esquemas para metadados como, o MARCXML e outros.

O período de transição de uma tecnologia para outra poderia ser breve se dependesse somente das tecnologias empregadas, mas geralmente esse período de transição é longo, pois, não depende somente das tecnologias, mas também e principalmente, da aceitação dessa nova tecnologia e de sua filosofia assim como a ruptura de paradigmas por parte dos profissionais de diversas áreas envolvidos na inovação tecnológica.

Steele (1983) citado por Jantz (2002, p. 68, tradução nossa) descreveu

o “desafio da inovação” como um processo que tem muitas barreiras. Para negociar com êxito a revolução tecnológica nas bibliotecas não somente necessitamos de novas idéias e um ambiente estimulante em que possam desenvolver-se, mas também gente que criam os novos produtos e que se encarreguem da difícil tarefa de desenvolvê-los. Estas pessoas são os defensores das novas idéias e necessitam de muito apoio para prosperar num

ambiente imerso na tradição. Sem a ajuda dessas pessoas não terão lugar as inovações.

Segundo Jantz (2002, p.63, tradução nossa) “A inovação é uma parte interessante do comportamento humano e tem aspectos misteriosos e desejáveis como a criatividade”.

Zairi (1992) citado por Jantz (2002, p.68, tradução nossa) oferece uma excelente definição de inovação tecnológica:

A inovação tecnológica é o processo pelo qual a indústria cria novos e melhores produtos e processos. Incluem atividades que vão desde a geração de uma idéia sobre um novo ou melhor produto, processo ou serviço, sua investigação, desenvolvimento e comercialização até a sua distribuição no mercado.

O que os profissionais da informação podem e devem fazer frente às inovações tecnológicas é aperfeiçoarem-se aprendendo sobre elas e conhecendo a fundo todos os aspectos relacionados com a inovação e a colaboração (JANTZ, 2002). Para Jantz (2002 p. 68, tradução nossa) “a inovação nos sugeri que devemos buscar paradigmas completamente novos para oferecer nossos serviços de informação e entender que, o que desconhecemos poderá ser mais importante do que basearmos no que já conhecemos”.

Grover, (1995 p. 8) nos diz que “devemos ser agentes que induzem mudanças ao invés de apenas reagirmos a pressões vindouras de afora”.

Frente a este novo espaço informacional, o profissional da informação se vê diante de uma nova direção na gestão da informação. Mas como diz Alvarenga (2001, p.14) “mudam-se os meios, sofisticam-se os instrumentos e surgem nomes novos para designar coisas velhas. Entretanto, a essência das coisas permanece”.

E, toda essa inovação tecnológica proporcionada pela XML e tecnologias associadas a ela, vem dar condições para se tornar realidade sonhos antigos, profetizados há

séculos e ainda não concretizados de fato, como a grande biblioteca universal, por exemplo, o controle bibliográfico universal, a biblioteca digital, dentre outros.

Mediante tudo isso, o formato de intercâmbio MARC21 é imprescindível para a catalogação cooperativa?

Através desta revisão de literatura observamos que o MARC foi o primeiro passo louvável para a consolidação da catalogação cooperativa em ambiente automatizado, e atualmente ele é utilizado por bibliotecas espalhadas por todo o mundo, formando redes cooperativas de dados catalográficos e bibliográficos, chegando a bilhões de registros disponíveis. No entanto, devido aos desenvolvimentos tecnológicos e ao uso intenso da Internet, torna-se necessário adaptá-lo a novas tecnologias emergentes, no caso a XML, que por sinal isso já vem ocorrendo, visto que esta é uma linguagem de marcação atual e que atende as necessidades informacionais da Web.

Desta forma, a utilização do formato MARC21 juntamente com tecnologias associadas a ele como a XML e outras, efetivaria uma catalogação cooperativa universal desde que haja uma política de adoção e investimentos nesse novo método e filosofia por parte dos profissionais e instituições.

Assim, o MARC21 em ambiente XML permitirá a interoperabilidade entre diversas linguagens e estruturas de metadados, facilitando o intercâmbio de dados entre diversas instituições, independente do sistema ou plataforma utilizados, além de proporcionar uma melhor interação entre humanos e máquinas devido ao esquema semântico da XML. E o MARC21, que até o momento representou eficientemente apenas registros catalográficos e bibliográficos, ou seja, registros tradicionais, em ambiente XML será capaz de representar registros que o MARC21 sozinho não suportaria como, textos completos, sons, gráficos e outros, representando tanto os dados primários (o próprio documento) como os dados secundários (metadados).

Podemos, através desta revisão de literatura apontar desenvolvimentos tecnológicos que vem ocorrendo na área da Biblioteconomia, e em especial na área da representação descritiva como, o **XMLMARC** desenvolvido pela Medlane, o **MARCXML** pela Library of Congress, entre outros.

Inclusive os grandes softwares de gerenciamento de bibliotecas utilizados por unidades de informação espalhadas pelo mundo inteiro já estão implementando uma camada XML em suas estruturas, como o **ALEPH500** por exemplo, buscando assim, facilitar o intercâmbio de dados entre as instituições por meio da interoperabilidade entre várias estruturas de metadados e sistemas heterogêneos utilizados (Informação verbal)⁸.

Portanto, torna-se necessário que os profissionais da informação de hoje não só aceite as mudanças mas sejam agentes destas, rompendo paradigmas, incorporando as mudanças, as inovações e as diversidades, visto que as necessidades sociais vigentes exigem isso.

Durante o desenvolvimento desta pesquisa pude vislumbrar diversos temas relacionados ao assunto aqui estudado e que poderiam servir para futuros trabalhos. A título de sugestão:

- O MARC em ambiente XML;
- O uso da XML nas Bibliotecas digitais;
- A coexistência de formatos para metadados nas bibliotecas como MARC21, MARCXML e outros esquemas para metadados;
- As Competências do profissional da informação frente às inovações tecnológicas;
- Metadados e o futuro da catalogação;

⁸ FONSECA, Marcelo Valle da. (Gerente da Ex Libris no Brasil). Palestra sobre: **Implantação de sistema de automação de biblioteca: Sistema ALEPH/2 – Metabuscaadores de informação estruturada: Sistema Metalib** - realizada em 25 de Novembro de 2003 no Anfiteatro II da Faculdade de Filosofia e Ciências – UNESP – Campus de Marília.

- Metadados e o Controle de Autoridades;
- Catalogadores na mediação e criação de sistemas de metadados;
- A criação de metadados no ato do desenvolvimento e publicação em arquivo eletrônico de um documento digital e a antiga proposta da catalogação na fonte.
- Catalogação de recursos da Internet com a atuação mais marcante do profissional bibliotecário;
- Parcerias entre bibliotecários, cientistas da informação e profissionais de outras áreas para desenvolver métodos para a descrição, organização e recuperação de objetos digitais na Internet;
- A utilidade do AACR2 para catalogar recursos da Internet.

Nesse sentido esperamos que este trabalho possa contribuir como uma atualização na forma de revisão de literatura para a área da Biblioteconomia e que venha a suscitar novas discussões sobre o assunto para assim dar continuidade a um tema tão atual, polêmico e imprescindível para a área.

REFERÊNCIAS

- ABITEBOUL, S. ; BUNEMAN P. ; SUCIU, D. **Gerenciando dados na Web**. Rio de Janeiro: Campus, 2000, 251 p. Tradução de: Mônica Córdia.
- ALVARENGA, L. A teoria do conceito revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. **Datagramazero – Revista de Ciência da Informação**, v.2, n.6, dez/2001. Disponível em: <http://www.dgzero.org/dez01/Art_05.htm>. Acesso em: 29 jan. 2003.
- AMARAL, V. **Desafios do trabalho em rede. 2002**. Disponível em <http://www.rits.org.br/redes_teste/rd_tmtes_dez2002.cfm>. Acesso em: 17 dez. 2002
- ARNOLD, P. ; WHITE, P. **A era da automação**. Rio de Janeiro: Lidador, 1965, 206 p. Tradução de: Hersílio Soares.
- BALBY, C. N. Formatos de intercâmbio de registros bibliográficos: conceitos básicos. **Cadernos da F.F.C.** Marília, v.4, n.1, 1995. p. 29-35.
- BAX, M. P. Introdução às linguagens de marcas. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 32- 38, jan./abr. 2001.
- BAX, M. P. **XML – Introdução e impactos nas bibliotecas**. 2002. Slides. Disponível em: <www.bax.com.br>. Acesso em: 20 nov. 2002.
- BERGHEL, H. The cost of having analog executives in a digital world. **Communications of the ACM**, v. 42, n. 11, p. 1115, 1999.
- BROOKSHEAR, J. G. **Ciência da Computação: uma visão abrangente**. 5. ed. Porto Alegre: Bookman, 2000. 499p.
- BRUN, R. E. **XML en la descripción de recursos**. Sevilha, 2002. Disponível em: <<http://www.forpas.us.es/aula/xml/doc/09.XML%20en%20la%20descripci%C3%B3n%20de%20recursos%20MARC.ppt>>. Acesso em: 20 nov. 2002.
- CAMPELLO, B. S.; MAGALHÃES, M. H. A. **Introdução ao controle bibliográfico**. Brasília: Brinquet de Lemos/Livros, 1997, 110 p.
- CARVALHO, F. Ciência aguarda dilúvio de informação. **O Estado de São Paulo**, 25 mar.1990, p. 24.
- CASTRO, E. **XML para a World Wide Web**. Rio de Janeiro: Campus, 2001. 269 p. (Visual quickstart guide). Tradução de Hugo de Souza Melo.
- CIANCONI, R. B. Banco de dados de acesso público. **Ciência da Informação**. Brasília, v.16, n.1, p. 53-59, jan./jun. 1987.

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.

CUNHA, M. B. **Base de dados e bibliotecas brasileiras**. Brasília: ABDF, 1984, 224 p.

EL-SHERBINI, M. Metadata and the future of cataloging. **Library Computing**, [S.l.], v.19, n.3/4, p.180-191, 2000.

FERREIRA, A. B. H. **Minidicionário da língua portuguesa**. 3. edição revista e ampliada. Rio de Janeiro: Nova Fronteira, 1993.

FERREIRA, S. M. S. P. Introdução às redes eletrônicas de comunicação. **Ciência da Informação**, Brasília, v.23, n.2, p. 258-263, maio./ago. 1994.

FIANDER, D. Applying XML to the Bibliographic Description. **Cataloging & Classification Quartely**, New York, v.33, n.2, p.17-28, 2001.

FOSKETT, D. J. **Informática**. In: Ciência da informação ou informática? Rio de Janeiro: Calunga, 1980. p.9-51.

FURGERI, S. **Utilização de catálogos XML para o desenvolvimento do comércio eletrônico**. 1999. 92 f. Dissertação (Mestrado em Sistemas de Informação) - Pontifícia Universidade Católica de Campinas, Instituto de Informática, Campinas, 1999.

FURGERI, S. **Ensino didático da linguagem XML**. São Paulo: Érica, 2001. 278 p.

FURRIE, B. **O MARC bibliográfico: um guia introdutório; catalogação legível por computador**. Brasília : Thesaurus, 2000. 95 p. Tradução de: Beatriz Valadares Cendón, Sonia Burnier, Maria Helena Santos e Natália Guiné de Mello Carvalho.

GAZAN, R. "The wonder years" of XML. **Library Computing**, [S.l.], v.19, n.1/2, p.13-17, 2000.

GRÁCIO, J. C. A. **Metadados para descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002. 127 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2002.

GRAMIGNA, M. R. M. **Jogos de empresa e técnicas vivenciais**. São Paulo: Makron Books, 1995, 216 p.

GROVER, M. L. Proposta para uma política de formação de acervos cooperativos. **Cadernos da FFC**. Marília, v.4, n.1, p. 1-8, 1995.

GROGAN, D. **A prática do serviço de referência**. Brasília: Brinquet de Lemos/Livros, 1995, 196 p. Tradução de: Antonio Agenor Brinquet de Lemos.

GUINCHAT, C. ; MENO, M. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. corr. aum. Por Marie France Blanquet. Brasília: IBICT, 1994. 540 p. Tradução de: Miriam Vieira da Cunha.

HARTUNG, A. P. **O cooperativismo** : ao alcance de todos. 2. ed. Florianópolis : OCESC, 1996.

HEEMANN, V. **Avaliação ergonômica de interfaces de bases de dados por meio de Checklist especializado**. 1997. Dissertação (Mestrado em Engenharia da Produção) – Universidade Federal de Santa Catarina, Florianópolis, 1997. Disponível em: <<http://www.eps.ufsc.br/disserta97/heemann/>>. Acesso em: 25 Set. 2003.

IOANNIDES, D. XML schema languages : beyond DTD. **Library Hi Tech**, [S. l.], v,18, n.1, p.9-14, 2000.

ISO - INTERNATIONAL ORGANIZATION STANDARDIZATION. 2002. Disponível em: <<http://www.iso.ch/iso/en/aboutiso/introduction/whatisISO.html> >. Acesso em: 15 abr. 2003.

JANTZ, R. C. Innovaciones tecnológicas en la biblioteca: proyectos digitales que ofrecen nuevas oportunidades al bibliotecario y a la biblioteca. **Boletín de la Asociación Andaluza de Bibliotecarios**, [S. l.], n. 67, p. 61-69, Junio de 2002.

JOHNSON, B. C. XML and MARC: which is “right”?. **Cataloging & Classification Quarterly**, New York, v. 32, n.1, p. 81-89, 2001.

KHURSHID, Z. From MARC to MARC21 and beyond: some reflections on MARC and the Arabic language. **Library Hi tech**, [S. l.], v.20, n.3, p.370-377, 2002.

LAM, K. XML and global name access control. **OCLC Systems & Services**, [S. l.], v.18, n.2, p. 88-96, 2002.

LANCASTER, F.W. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 1991. 347 p.

LEHNUS, D. J. **Catálogo descritiva**. Brasília: VIPA, 1971, 116 p. (Visão da informática pura e aplicada, v.2). Tradução, adaptação e ampliação de Abner Lellis Corrêa Vicentini.

LEITÃO, S. P; ROSSI, L. P. Natureza da mudança nas organizações: uma introdução crítica. **Revista Arché Interdisciplinar**, v.10, n.28, p.13-38, 2000.

LIBERTY, J. ; KRALEY, M. **Aprendendo a desenvolver documentos XML para a Web**. São Paulo: Makron Books, 2001.274 p. Tradução de: Flávia Cruz.

MARCONDES, C. H. Representação e economia da informação. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 61-70, jan./ abr. 2001.

MARTÍNEZ GONZÁLEZ, M. M. Extended Markup Language (XML) : Una solución para modelar documentos y sus interrelaciones basada en la semántica de la información. **Scire: Representación y Organización del Conocimiento**, Zaragoza, v.6, n.2, p.121-151, jul./dic. 2000.

MERCADANTE, L. Integração a redes de catalogação cooperativa: UNICAMP x BIBLIODATA/CALCO. **Cadernos da F.F.C**. Marília, v.4, n.1, 1995. p. 24-28.

- MEY, E. S. A. **Introdução à catalogação**. Brasília: Brinquet de Lemos/Livros, 1995, 123 p.
- MILSTEAD, J. ; FELDMAN, S. Metadata: Cataloging by Any Other Name. **Online**, Jan. 1999. Disponível em: <<http://www.onlineinc.com/online/OL1999/milstead1.html>>. Acesso em: 10 set. 2002.
- MUELLER, S. P. M. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Marguerite (Org.). **Fontes de Informação para Pesquisadores e Profissionais**. Belo Horizonte, Ed. UFMG, 2000, p. 21 – 34.
- NISO** - NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.2-1994**: information interchange format. United States of America, 1994. 6 p. Disponível em: <<http://www.niso.org/standards/resources/Z39.2.PDF>>. Acesso em: 15 abr. 2003.
- NEEDLEMAN, M. et al. The NISO circulation interchange protocol (NCIP) – an XML based Standard. **Library Hi Tech**, [S. l.], v.19, n.3, p. 223-230, 2001.
- NORONHA, D. P. ; FERREIRA, S. M. S. P. Revisões de Literatura. In: CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Marguerite (Org.). **Fontes de Informação para Pesquisadores e Profissionais**. Belo Horizonte, Ed. UFMG, 2000, p.191 – 198.
- ORTIZ-REPISO JIMÉNEZ, V. Qué enseñamos después Del MARC?. **Organización del Conocimiento en Sistemas de Información y Documentación**. Zaragoza, v.3, p. 217-225, 1999. Actas del III Encuentro de ISKO-España Getafe, 19 al 21 de noviembre de 1997.
- QIN, J. Representation and Organization of Information in the Web Space: From MARC to XML. **Informing Science**, [S. l.], v.3, n.2, p. 83-87, 2000.
- RAMALHO, J. A. **XML teoria e prática**. São Paulo: Berkeley Brasil, 2002, 146 p.
- RAY, E. T. **Aprendendo XML**. Rio de Janeiro: Campus, 2001. 372 p. Tradução de: Daniel Vieira.
- RHYNO, A. XML and relational databases : Uses and opportunities for libraries. **OCLC Systems & Services**, [S. l.], v. 18, n.2, p.97-103, 2002.
- ROSSETO, M. **Uso do protocolo Z39.50 para recuperação de informação em redes eletrônicas**, 1997. Disponível em:<<http://www.internic.net/z3950/z3950.html>>. Acesso em 10 abr. 2002.
- ROWLEY, J. **Informática para bibliotecas**. Brasília: Brinquet de Lemos/Livros, 1994, 307 p. Tradução de: Antonio Agenor Brinquet de Lemos.
- ROWLEY, J. **A biblioteca eletrônica**. 2. edição de informática para bibliotecas. Brasília: Brinquet de Lemos/Livros, 2002, 399 p.
- SAMBAQUY, L.Q. Catalogação Cooperativa e Catalogação Centralizada. **Revista do Serviço Público**. ago., 1951. p. 36-39.
- SANTOS, P. L. V. A. C. **Comunicação**. Marília, Unesp, 2002. 10f. Notas de aula.

SHAPIRO, C. ; VARIAN, H. R. **A economia da informação**: como os princípios econômicos se aplicam a era da Internet. 5ª edição, Rio de Janeiro: Campus, 1999. Tradução de: Information rules por Ricardo Inofosa. 397 p.

SILVA, A. F. G. Marketing em unidades de informação: revisão crítica. **Revista de Biblioteconomia de Brasília**, v. 23/24, n.1, p 5-24, especial 1999/2000.

SOUZA, M. I. F. ; VENDRUSCULO, L. G. ; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica : utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v. 29, n.1, p. 93-102, jan./ abr. 2000.

SOUZA, T. B. Catalogação Cooperativa na Rede Bibliodata/CALCO: a repetitividade de registros bibliográficos no catálogo coletivo. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECONOMIA “DR. PAULO TARCÍSIO MAYRINK”, 3., 1999, Marília. **Anais...** Marília: UNESP, 1999, p. 247-258.

SOUZA, T. B. ; CATARINO, M. E. ; SANTOS, P. C. Metadados: catalogando dados na Internet. **Transinformação**, v.9, n.2, mai/ago, 1997.

STEELE, L. Managers' misconceptions about technology. **Harvard Business Review**, p. 133-141, Nov./Dez., 1983.

STEWART, M. **MARC harmonization update**: part I. background to MARC21. In: CLA'99, 1999, Toronto. Disponível em:
<<http://www.fis.utoronto.ca/people/affiliated/tsig/stewart.html>>. Acesso em: 18 dez. 2001.

TARAPANOFF, K. Cenários para o planejamento estratégico de unidades de informação voltados para a qualidade. **Documento datilografado**. 1996.

TENNANT, R. MARC Must Die. **Library Journal**, 2002a. Disponível em :
<<http://libraryjournal.reviewsnews.com/index.asp?layout=article&articleid=CA250046&display=Digital+LibrariesNews&industry=Digital+Libraries&industryid=3760&verticalid=151>>. Acesso em: 20 set. 2002

TENNANT, R. MARC Exit Strategies. **Library Journal**, 2002b. Disponível em:
<<http://libraryjournal.reviewsnews.com/index.asp?layout=article&articleid=CA256611&display=Digital+LibrariesNews&industry=Digital+Libraries&industryid=3760&verticalid=151>>. Acesso em: 20 set. 2002.

W3C WORLD WIDE WEB CONSORTIUM. 1994 – 2002. Disponível em:
<<http://www.w3c.org/consortium>> 20 de maio 2002.

ZAIRI, M. Managing usersupplied interactions: management of R&D activity. **Management Decision**, v.30, n. 8, p. 49-57, 1992.

ZOCCO, L. P. **Estudo comparado dos Formatos de Intercâmbios bibliográficos e catalográficos**. Marília: UNESP, 1993.

BIBLIOGRAFIA CONSULTADA

AHRONHEIM, J. Exploding out of the MARC Box : Building New Roles for Cataloging Departments. **Cataloging & Classification Quartely**, New York, v.30, n.2/3, p. 217-225, 2000.

ALBANO, A. et. al. A typed text retrieval query language for XML documents. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 467-488, abr. 2002.

ALMEIDA, M. B. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. **Ciência da Informação**, Brasília, v.31, n.2, p.5-13, maio/ago. 2002.

ALVES, R. C. V. **Análise dos padrões de descrição das informações para organização de documentos eletrônicos**: um estudo sobre metadados. Marília: UNESP, 2002. 21f. Relatório parcial de bolsa de Iniciação Científica CNPq.

ANDERSON, G. Symmetry and Extrapolation: Passion and Precision – Cooperative Cataloging at the Beginning of the 21st Century. **Cataloging & Classification Quartely**, New York, v.17, n.3/4, p.39-73, 1993.

ANGÓS ULLATE, J. M. ; SALVADOR OLIVÁN, J. A. ; FERNÁNDEZ RUÍZ, M. J. Metadatos en los documentos HTML: una ayuda para la recuperación de información. **Scire: Representación y Organización del Conocimiento**, Zaragoza, v.5, n.2, p.47-61, jul./dic. 1999.

ARARIPE, F. M. A. ; MOURA, R. C. O. Cooperativismo no mercado informacional do Mercosul: uma alternativa. **Informação & Informação**, Londrina, v.5, n.2, p.103-110, jul./dez. 2000.

AYRES, F. H. Authority Control Simply Does Not Work. **Cataloging & Classification Quartely**, New York, v.32, n.2, p.49-59, 2001.

AZAGURY, A. et. al. A novel navigation paradigm for XML repositories. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 515-525, abr. 2002.

BAPTISTA, A. A. ; MACHADO, A. B. Um gato preto num quarto escuro – falando sobre metadados. **Revista de Biblioteconomia de Brasília**, v.25, n.1, p.77-90, jan.jun. 2001.

BARBOSA, A. P. **Novos rumos da catalogação**. Organização, revisão e atualização de Elza Lima e Silva Maia. Rio de Janeiro: BNG/Brasilart, 1978. 246 p. (Coleção Biblioteconomia, documentação, ciência da informação).

BARONE, F. A revisão de literatura em biblioteconomia e documentação: uma prática pedagógica. **Revista Brasileira de Biblioteconomia e Documentação**. São Paulo, v.23, n.1/4, p.136-146, jan.dez. 1990.

- BARRETO, A. A. Os destinos da Ciência da informação: entre o cristal e a chama. **DataGramZero- Revista de Ciência da Informação**, n. zero, dez. 1999. Disponível em: <http://www.dgz.org.br/dez99/Art_03.htm>. Acesso em 10 out 2002.
- BAX, M. P. As bibliotecas na Web e vice-versa. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.3, n.1, p.5-20, jan./jun. 1998.
- BORDANA VE, J. E. D. **O que é comunicação**. 6.ed. São Paulo: Brasiliense, 1985. 104 p.
- BOWEN, Jennifer B. Creating a culture of cooperation. **Cataloging & Classification Quarterly**, New York, v.26, n.3, p.73-85, 1998.
- BRASIL. Ministério da Ciência e Tecnologia. **Sociedade da informação no Brasil**: livro verde. Brasília, DF, 2000. 195p.
- BROWN, D. R. Consórcios e redes nas bibliotecas acadêmicas dos EUA. **Transinformação**, v.10, n.1, p.33-61, jan./abr, 1998.
- BUTTERFIELD, K. L. Cataloger's and the creation of metadata systems: a collaborative vision at the University of Michigan. **OCLC Internet Cataloging Project Colloquium**, San Antonio, Texas, January, 19, 1996. Disponível em: <<http://www.oclc.org/oclc/man/colloq/butter.htm>>. Acesso em: 07 maio 2003.
- CALVO, A. M. Structuring biographical data in EAD with the Nomen DTD. **OCLC Systems & Services**, [S. l.], v.17, n.4, p.187-199, 2001.
- CAPLAN, P. "You call it corn, we call it syntax-independent metadata for document-like objects." **The Public-Access Computer Systems Review**, v.6, n. 4, 1995. Disponível em: <<http://info.lib.uh.edu/pr/v6/n4/capl6n4.html>>. Acesso em: 07 maio 2003.
- CHAN, A. T. S. et. al. A survey in indexing and searching XML documents. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 415-437, abr. 2002.
- CHANG, M. An electronic finding aid using extensible markup language (XML) and encoded archival description (EAD). **Library Hi Tech**, [S. l.], v.18, n.1, p.15-27, 2000.
- CÔRTE, Adelaide Ramos e; ALMEIDA, Iêda Muniz de (Coord.). **Avaliação de software para bibliotecas**. São Paulo: Polis, APB, 2000. 108p. (coleção palavra-chave, 11).
- COYLE, K. Open Source, Open Standards. **Information Technology and Libraries**, [S. l.], v.--, n.--, p.33- 36, Mar. 2002.
- DELGADO GÓMEZ, A. Aplicaciones de descripción documental y recuperación de la información en el entorno del fin de siglo y en marco de las bibliotecas públicas. **Scire: Representación y Organización del Conocimiento**, Zaragoza, v.5, n.2, p.63-73, jul./dic. 1999.
- DIAS, E. J. W. OCLC: um novo conceito em cooperação bibliotecária. In: 9º Congresso Brasileiro & V Jornada Sul - Rio-Grandense de Biblioteconomia e Documentação, 1977, Porto Alegre. **Anais...** Porto Alegre, v1, 1977. p. 38-45. 2v.

FERNÁNDEZ-MOLINA, J. C. Uso de esquemas de clasificación para mejorar las prestaciones de visualización (browsing) de los catálogos en línea. **Organización del Conocimiento en Sistemas de Información y Documentación**. Zaragoza, v.3, p. 295-305, 1999.

FIGUEIREDO, N. Da importância dos artigos de revisão da literatura. **Revista Brasileira de Biblioteconomia e documentação**. São Paulo, v.23, n.1/4, p.131-135, jan./dez. 1990.

FLAMINO, A. N. et al. **Catálogo Cooperativa**. Marília, UNESP, 2001. 31f. Seminário apresentado na disciplina Catálogo Automatizada ministrada pela Profª. Drª. Plácida Leopoldina Ventura Amorim da Costa Santos.

FLAMINO, A. N. et al. **Evolução da Internet**. Marília, UNESP, 2000. 30f. Seminário apresentado na disciplina Introdução à Ciência da Computação ministrada pela Profª Drª. Silvana Aparecida Borsetti Gregório Vidotti.

FURNIVAL, A. C. Redes para cooperação internacional: uma janela de oportunidade para os países recém industrializados?. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.3, n.1, p.115-123, jul./dez. 1998.

GARCÍA MARTÍNEZ, A. M. Definición y estilo de los objetos de información digitales y metadatos para la descripción. **Boletín de la Asociación Andaluza de Bibliotecarios**, [S. l.], n. 63, p. 23-47, Junio de 2001.

GARDNER, J. R. Information architecture planning with XML. **Library Hi Tech**, [S. l.], v.19, n.3, p. 231-241, 2001.

GUSMÃO, A. O. M. **Avaliação da qualidade e determinantes de desempenho do Aleph 500 em bibliotecas universitárias Brasileiras**. 2001. 213 f. Dissertação (Mestrado em Ciência da Informação) Universidade Federal da Paraíba, João Pessoa.

HANCOCK, B. ; GIARLO, M. F. Moving to XML : Latin texts XML Conversion Project at the Center for Electronic texts in the Humanities. **Library Hi Tech**, [S. l.], v. 19, n.3, p.257-263, 2001.

HIGHSMITH, A. et al. MARC It Your Way : MARC. P.m. **Information Technology and Libraries**, [S. l.], p.19-25, Mar. 2002.

HOPKINSON, A. O “Common Communication Format” desenvolvido pela UNESCO. **Ciência da Informação**, Brasília, v. 14, n.1, p.51-54, jan./jun. 1985.

IKEMATU, R. S. Gestão de metadados: sua evolução na tecnologia da informação. **DataGramZero- Revista de Ciência da Informação**, v.2, n.6, dez. 2001. Disponível em: <http://www.dgzero.org/dez01/Art_02.htm>. Acesso em: 29 jan. 2002.

JOACHIM, M. D. Cooperative Cataloging of Microform Sets. **Cataloging & Classification Quarterly**, New York, v.17, n.3/4, p.105-126, 1993.

JULICH, S. ; HIRST, D. ; THOMPSON, B. A case study of ILS migration: Aleph500 at the University of Iowa. **Library Hi Tech**, v.21, n.1, p.44-55, 2003.

KIM, H. ; CHOI, C. XML : how it will be applied to digital library systems. **The Electronic Library**, [S. l.], v.18, n.3, p.183-189, 2000.

KUSHMERICK, N; CHINENYANGA, T. T. An expressive and efficient language for XML information retrieval. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 438-453, abr. 2002.

LE COADIC, Y. **A ciência da informação**. Brasília: Briquet de Lemos/Livros, 1996. 119p. Tradução de Maria Yêda F. S. de Filgueiras Gomes.

LEWIS, J. XML : an introduction. **OCLC Systems & Services** , [S.l.], v.14, n.1, p.51-52, 1998.

LÓPEZ-HUERTAS PÉREZ, M. J. Lãs relaciones semânticas: su papel em la materialización de la representación y organización del conocimiento. **Organización del Conocimiento en Sistemas de Información y Documentación**. Zaragoza, v.3, p. 57-67, 1999.

MARASCO, L. C. Avaliação e seleção de software para automação de centros de documentação e bibliotecas. **Cadernos da FFC. Marília**, v.4, n.1, p. 40-49, 1995.

MARC STANDARDS. Disponível em: <<http://lcweb.loc.gov/marc>>. Acesso em: 20 nov. 2001.

MARC21: formato condensado para dados bibliográficos. Tradução e adaptação: Margarida M. Ferreira. 2. ed. Marília: Unesp – Marília – Publicações, 2002. v.1. 247 p. Tradução e adaptação de: MARC 21 concise for bibliographic data da Network development and MARC Standards Office – Library of Congress – USA.

MARCONDES, C. H. ; SAYÃO, L. F. Documentos digitais e novas formas de cooperação entre sistemas de informação em C & T. **Ciência da Informação**, Brasília, v.31, n.3, p.42-54, set./dez. 2002.

MARXML. Disponível em:<<http://www.loc.gov/standards/marxml>>. Acesso em: 20 maio 2002 e 29 ago. 2002.

McGARRY, K. **O contexto dinâmico da informação**: uma análise introdutória. Tradução de Helena Vilar de Lemos. Brasília: Brinquet de Lemos/Livros, 1999. 206 p.

MEDEIROS, N. Peering over the fortress walls: The metadata invasion begins. **OCLC Systems & Services**, [S. l.], v.17, n..4, p. 154-156, 2001.

MERLO VEJA, J. A. Organizaciones de normalización en Internet. **Revista Española de Documentación Científica**, [S. l.], v.23, n.2, p. 327-340, jul./sept. 2000.

MEUSS, H.; SCHLIEDER, T. Querying and ranking XML documents. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 489-503, abr. 2002.

MILSTEAD, J.; FELDMAN, S. Metadata projects and standards. **Online**, Jan. 1999. Disponível em: <<http://www.onlineinc.com/online/OL1999/milstead1.html>>. Acesso em: 10 set. 2002.

MORGAN, E. L. Possibilities for Open Source Software in Libraries. **Information Technology and Libraries**, [S. l.], p. 12-15, Mar. 2002.

MOURA, A. M. S. ; CORREIA, A. E. G. C. Hipertexto para o ensino da representação descritiva: uma experiência de modelagem. **Informação & Sociedade**, João Pessoa, v.10, n.1, p.73-80, 2000.

NAVARRO, G.; BAEZA-YATES, R. XML and proximal nodes. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 504-514, abr. 2002.

NEEDLEMAN, M. Z39.50 – a review, analysis and some thoughts on the future. **Library Hi Tech**, [S. l.], v.18, n.2, p.158-165, 2000.

PEDRAZA GRACIA, M. J. Aproximación a la catalogación de documentos electrónicos de acceso remoto. **Scire: Representación y Organización del Conocimiento**, Zaragoza, v.3, n.1, p. 99-125, ene./jun. 1997.

PEREIRA, A. M. **Catálogo e o uso de tecnologias**. 1996. 80f. Trabalho de conclusão de curso (Bacharel em Biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

PRADO, N. S. Utilizando o campo 856 do MARC para disponibilizar texto integral da produção docente da UDESC na Internet. **Transinformação**, Campinas, v.10, n.2, maio/ago. 1998. Disponível em:
<<http://www.puccamp.br/~biblio/transinformacao/old/vol10n2/pag104.html>>.
Acesso em: 23 abr. 2003.

RIGGS, K. R. XML and free text. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 526-528, abr. 2002.

ROJO, A. S. ; VEGA, J. A. M. Organizaciones de normalización en Internet. **Revista Española de Documentación Científica**, v.23, n.2, p.327-340, jul./sept. 2000.

ROSA PIÑERO, A. ; SENSO, J. A. El concepto de metadato. Algo más que descripción de recursos electrónicos. **Ciência da Informação**, Brasília, v.32, n.2, p. 95-106, maio/ago. 2003.

SANDBERG-FOX, A. M. Report on the Library of Congress Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web, Held Nov. 15-17, 2000, in Washington, DC. **Cataloging & Classification Quarterly**, New York, v.32, n.2, p.35-48, 2001.

SANT'ANA, R. C. G. ; SANTOS, P. L. V. A. C. Transferência da Informação: análise para valoração de unidades de conhecimento. **Datagrama zero – Revista de Ciência da Informação**, v.3, n.2, abril. 2002. Disponível em:
<http://www.dzg.org.br/abr02/F_I_art.htm>. Acesso em: 08 maio 2003.

SANTOS, P. L. V. A. C. **Catálogo Automatizada**. Marília, UNESP, 2001. 18f. Notas de aula.

SANTOS, P. L. V. A. C. Representação descritiva e seu ensino nos cursos de graduação em biblioteconomia. **Cadernos da FFC. Marília**, v.4, n.1, p.36-39, 1995.

SANTOS, P. L. V. A. C. ; MARTINEZ, V. C. A rede e o conhecimento. **Informação & Informação**, Londrina, v.5, n.2, p.111-124, jul.dez. 2000.

SEREBRENİK, A. et. al. EquiX- A search and query language for XML. **Journal of the American Society for Information Science and Technology**. v.53, n.6, p. 454-466, abr. 2002.

SIQUEIRA, M.A. **XML na Ciência da Informação**: uma análise do MARC21. 2003. 134 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

STEINHAGEN, E. ; MOYNAHAN, S. Catalogers must change! Surviving between the rock and the hard place. **Cataloging & Classification Quartely**, New York, v.26, n.3, p. 3-20, 1998.

SWANEKAMP, J. The changing cataloging culture: what do we mean when we talk about cataloger values?. **Cataloging & Classification Quartely**, New York, v.26, n.3, p.51-55, 1998.

TAYLOR, A. G. ; CLEMSON, P. Access to networked documents: Catalogs? Search engines? Both?. **OCLC Internet Cataloging Project Colloquium**, San Antonio, Texas, January, 19, 1996. Disponível em: <<http://www.oclc.org/oclc/man/colloq/taylor.htm#fig>>. Acesso em: 07 maio 2003.

TCP/IP. Disponível em: <<http://sites.uol.com.br/luizjrcfox/tcp.htm>>. Acesso em: 20 maio. 2002.

TILLET, B. Catalog It once for All : A History of Cooperative Cataloging in the United States Prior to 1967 (Before MARC). **Cataloging & Classification Quartely**, New York, v.17, n.3/4, p.3-38, 1993.

TRISKA, R. ; CAFÉ, L. Arquivos Abertos: subprojeto da Biblioteca Digital Brasileira. **Ciência da Informação**, Brasília, v.30, n.3, p.92-96, set./dez. 2001.

VALENTIM, M. L. P. Estrutura de Bases de dados: modelos de metadados e a qualidade de resposta. **Transinformação**, Campinas, v. 13, n. 1, p. 67-80, jan./jun. 2001.

VASCONCELOS, P. A. G. BIBLIODATA/CALCO: informação bibliográfica para o desenvolvimento. **Ciência da Informação**, Brasília, v.25, n.3, p.1-6, 1996.

VITA, S. H. PREMARC: a history and status report. **Cataloging & Classification Quartely**, New York, v.14, n.314, 1992.

W3C WORLD WIDE WEB CONSORTIUM. 1994 – 2002. Disponível em: <<http://www.w3c.org/consortium>>. Acesso em: 20 de maio. 2002.

WUSTEMAN, J. XML to desktop. **Library Hi Tech**, v.21, n.2, p.238-245, 2003.

XML e ACCESS 2002. Disponível em: <<http://office.microsoft.com/brasil/assitance/2002/articles/acAboutXmlDataAndAccess>>. Acesso em: 20 maio. 2002.

XMLMARC. Disponível em: <<http://xmlmarc.stanford.edu>>. Acesso em: 20 maio 2002.

XU, A. Accessing information on the Internet: feasibility study of USMARC format and AACR2. **OCLC Internet Cataloging Project Colloquium**, San Antonio, Texas, January 19, 1996. Disponível em:< <http://www.oclc.org/oclc/man/colloq/xu.htm>>. Acesso em: 07 maio 2003.

YACHNES, P. XMLcat : na XML-encoded online library catalog. **Library Computing**, [S. l.], v.19, n.1/2, p.59-67, 2000.