

Artigo
**Meta-Dados Aplicado ao Processo de
Desenvolvimento de Coleções**

Autores do Artigo

Marcos dos Santos,

Lucileide Andrade de Lima do Nascimento, Elias Oliveira

Departamento de Ciências da Informação

Universidade Federal do Espírito Santo

Campus de Goiabeiras, Av. Fernando Ferrari, s/n,

Cx Postal 5011, 29060-970 – Vitória, ES.

blue <http://www.inf.ufes.br/~elias>

elias@inf.ufes.br

Área do Trabalho

Uso da tecnologia e informação e comunicação - TIC;

Meta-Dados Aplicado ao Processo de Desenvolvimento de Coleções

Marcos dos Santos,
Lucileide Andrade de Lima do Nascimento e Elias Oliveira
Departamento de Ciências da Informação
Universidade Federal do Espírito Santo
Campus de Goiabeiras, Av. Fernando Ferrari, s/n,
Cx Postal 5011, 29060-970 – Vitória, ES.
blue <http://www.inf..ufes.br/~elias>
elias@inf.ufes.br

Resumo: Apresenta a utilização de metadados de planos de disciplinas como mais um instrumento de operacionalização para geração de indicadores de demanda no desenvolvimento de coleções. Através do uso de meta-estruturas descritas na linguagem XML, representa tanto as informações contidas nos programas de disciplinas do Departamento de Ciências da Informação da UFES, como também o acervo da biblioteca. Com esta base de dados formada, utiliza uma outra ferramenta associada a tecnologia XML para o processamento automático destas informações: a linguagem de processamento e transformação de dados XSLT. Com a utilização destas tecnologias foi possível a criação automática de uma página eletrônica apontando as necessidades demandadas no acervo pelas disciplinas do departamento à medida que as mesmas são introduzidas na base de dados. Além disso, através de tecnologias associadas à linguagem de metadados XML, é possível também acessar base de dados de livrarias, como por exemplo a www.amazon.com e obter de forma automática os preços das obras a serem incorporadas ao acervo.

Palavras-chave: Metadados, XML, Processamento da Informação, Desenvolvimento de Coleções.

1 Introdução

O papel de provisão da informação desempenhado por uma biblioteca universitária assenta-se na capacidade que esta possui em selecionar e disponibilizar o material bibliográfico pertinente às necessidades informacionais de seus usuários. Para desempenhar esse papel a própria biblioteca, por sua vez, demanda constantemente colher informações em sua própria coleção além de gerenciar os diversos formatos e acompanhar os novos produtos informacionais no mercado, de modo a contrabalancear o seu coleção sem prejuízos à sua clientela e aos cofres da instituição onde está inserida. Nesse sentido a atividade de avaliação de coleções é um componente chave nesse processo, visto que o mesmo relaciona-se diretamente a vários elementos que caracterizam a coleção, como seu conteúdo (cobertura, autoridade, atualidade), seu grau de adequação ao usuário (idioma, relevância) e aspectos adicionais relativos ao formato (aspectos físicos, custo), possibilitando dessa forma traçar vários perfis acerca dessa coleção, seja em relação às necessidades informacionais, seja em relação a conveniência e portabilidade de um formato pelos usuários, entre outros aspectos. Entretanto, apesar de sua importância, a avaliação de coleções não tem sido inserida às rotinas das

bibliotecas universitárias de forma contundente, se por um lado os bibliotecários parecem apresentar carências em relação ao conhecimento de metodologias de aplicação nessa atividade em uma unidade de informação, por outro lado tem-se o reflexo direto dessa atitude na qualidade do material informacional a ser disponibilizado aos usuários. Adiciona-se a isso o fator contingente reduzido de pessoal para o grande tamanho do coleção, em muitas bibliotecas.

Pensamos que para aliviar a falta de pessoal, o bibliotecário terá que cada vez mais, efetivamente, lançar mão de recursos tecnológicos que o auxiliem nessa atividade e outras que tenha que realizar, principalmente aquelas relacionadas ao controle e monitoramento dinâmico da coleção. A possibilidade do próprio profissional que atua na biblioteca através da sua criatividade, moldar uma ferramenta em seu domínio não é uma coisa exógena, mas deverá se tornar mais corriqueira nos próximos anos.

A construção de habilidades no efetivo uso de metalinguagens para representação de dados pode abrir muitas portas para soluções criativas em uma unidade de informação. Este trabalho procura mostrar que estas ferramentas, concretizadas na linguagem XML (*Extensible Markup Language*)(Ray 2003) e em outras tecnologias associadas, sendo utilizadas para a representação de programas de disciplinas de departamentos acadêmicos, podem fornecer ao gestor de coleção informações preciosas, e isso feito de forma automática.

Este trabalho está organizado da seguinte forma. Na Seção 2 percorremos, de forma breve, a literatura a respeito de alguns exemplos de como podemos obter informações para subsidiar a gestão de coleções em uma biblioteca universitária. Apresentamos, na Seção 3, uma breve introdução à linguagem de marcação XML, linguagem utilizada neste trabalho para construção de metadados para plano de disciplinas. Nesta seção apresentamos também algumas tecnologias de processamento de metadados associadas à linguagem XML. Apresentamos na Seção 4 a forma com a qual construímos os metadados e como também realizamos o processamento destes. Para validarmos nossa proposta, foi implementado um protótipo que será apresentado na Seção 5, onde apresentamos alguns dos resultados informacionais produzidos, de forma automática e diretamente para uma página de WEB. Finalmente, na Seção 6, apresentamos as conclusões deste trabalho e apontamos alguns possíveis desdobramentos para trabalhos futuros.

2 Revisão de Literatura

Na literatura existem algumas propostas para estratégias de tomada de decisão etapa de seleção de novos itens para inclusão ou complementação de coleção em uma biblioteca. Estas vão desde a consulta direta aos membros dos departamentos de uma universidade, até a utilização de um procedimento mais sofisticado de inferência (Danilowicz e Szarski 1981, Kao, Chang e Lin 2003).

No trabalho apresentado em (Kao, Chang e Lin 2003) a técnica de *Datamining* foi utilizada para a consideração dos itens a serem levados em conta no processo de aquisição de coleção. *Datamining* é um processo de se descobrir conhecimentos implícitos em uma grande base de dados. Esta técnica tem a capacidade de revelar relações, padrões e tendências que muitas vezes podem passar despercebidas em uma base de dados histórica (Kantardzic 2002).

O processo apresentado em (Kao, Chang e Lin 2003), batizado como ABAMDM (*Acquisition Budget Allocation Model via Data Mining*), consiste em obter, através dos dados históricos da circulação diária dos itens informacionais da coleção, o quanto os usuários de cada departamento se utilizam de itens específicos de determinada coleção.

Esta contabilização mostrará, do total de membros de um departamento **A**, por exemplo, o quanto o mesmo demanda apropriadamente por seu subconjunto do coleção geral da biblioteca. Através desta contabilização, o departamento **A** pode receber um peso maior do que o departamento **B** que sistematicamente utiliza mais itens de outras áreas. Desta forma, no momento de se tomar uma decisão sobre a distribuição de verba para a aquisição de novos itens, o departamento **A** receberia uma quantia maior da verba que o departamento **B**. É evidente que esta regra é altamente subjetiva, como bem observam seus autores. Entretanto, um ponto importante a se observar nesse procedimento é o não importúneo ao usuário na obtenção da informação. Todo o processo é realizado silenciosamente nas base de dados da biblioteca.

(Danilowicz e Szarski 1981), trabalhando com periódicos científicos, reúne três classes de estratégias comumente usadas na literatura para qualificação da importância desse tipo de coleção: 1) Pedir a opinião do usuário sobre uma dada relação de periódicos (*check list*); 2) Conferir a frequência do acesso, em um dado tempo, ao periódico na biblioteca; 3) Avaliar a quantidade de citações feitas aos artigos publicados na revista em outras mídias científicas.

Parece-nos, a partir de observações empíricas, que a opção de *check list* seja a mais utilizada no contexto brasileiro, seja via manual, seja via o próprio sistema de automação de biblioteca. Isto não somente para os periódicos, mas também para outros itens informacionais do coleção da biblioteca.

Os dois outros procedimentos, 2) e 3) acima, possibilitam uma avaliação mais precisa da real necessidade do item informacional no coleção. Ao contrário do primeiro que depende de uma consulta direta aos usuários, os dois últimos procedimentos dependem apenas de análises feitas na base de dados da biblioteca sobre os itens já existentes no coleção.

O autor em (Oliveira 2004) explora mais uma metodologia de se obter a informação de demanda de um departamento acadêmico. O processo se baseia no monitoramento das referências bibliográficas elencadas pelos próprios usuários da biblioteca em seus trabalhos científicos. Este monitoramento possibilitará uma avaliação quantitativa da importância do item requisitado e, além disso, a revelação de novos itens para a biblioteca. Muito embora o autor apresenta um exercício manual para comprovação de sua metodologia, tal processo poderia ser igualmente automatizado ao apresentado no presente trabalho.

Na tentativa de expressar formalmente um sistema ideal de serviço de informação, (Danilowicz e Szarski 1981) nos dão a seguinte definição. Seja **U** o conjunto dos usuários com interesse em periódicos científicos em um particular assunto ou área, e **J** o conjunto de todas os periódicos que fornecem informação de interesse dos usuários em **U** (Danilowicz e Szarski 1981). Desta forma, consideram um sistema de informação ideal se:

- a) Todos os artigos do conjunto **J** estão na conta de interesses dos usuários;
- b) O sistema provê serviço de acesso para todos os usuários em **U**;
- c) A informação prestada aos usuários é pertinente e completa.

Em outras palavras, um sistema de informação ideal é tal que provê a todos os usuários de cada segmento do conhecimento todos os artigos de interesse dos mesmos. Como tal sistema ideal não existe, os sistemas existentes são apenas uma aproximação do desejado. Diante disso, trabalhamos sempre com uma aproximação do que poderia ser ideal. Portanto, a proposta que aqui apresentamos não invalida as anteriormente

apresentadas, mas as complementa com o intuito de melhorar ainda mais tal aproximação.

Na próxima seção descrevemos brevemente a estrutura da linguagem XML, a qual será utilizada como ferramenta para construção de metadados de programas de disciplinas. Nos programas de disciplinas, os docentes além de explicitarem suas necessidades de fontes informacionais, também informam a quantidade de usuário para as mesmas. Infelizmente, por vezes, esta informação é negligenciada pelas administrações e pelos gestores da coleção.

3 A Linguagem XML e Utilitários

O formato MARC para representação e descrição de documentos, foi criado com o objetivo de termos uma linguagem *legível por máquina*. Portanto, o termo legível por máquina, ou computador, significa que um determinado documento descrito nessa linguagem pode ser lido, manipulado e interpretado mais facilmente por uma máquina.

O XML (*Extensible Markup Language*) se enquadra nesta categoria de linguagens. Desta forma, assim como estruturado no formato MARC, resguardada as devidas proporções, existe na linguagem XML *sinalizadores* antes de cada elemento representado. São os chamados *tags*, ou marcadores (Ray 2003). Estes marcadores, na verdade, servem de auxiliares para que o computador possa ler e interpretar os campos por eles demarcados no registro de um item.

Há varias décadas que se recorre à marcação para estruturar documentos. O SGML (*Standard Generalize Markup Language*) é a linguagem da qual muitas outras se originaram: HTML (*Hyper Text Markup Language*), XML, entre outras.

A linguagem HTML, na prática, tornou-se inviável para a representação da estrutura de documentos, uma vez que o seu desenvolvimento acabou se desviando do objetivo a que pretendia: ser uma alternativa a SGML, que é uma linguagem fortemente estruturadora de documentos. Ao invés disso, HTML transformou-se em uma linguagem apenas de formatação de documentos. Por outro lado, a SGML ainda constituía-se em uma linguagem complexa, ficando restrita a poucos especialistas. A resposta a SGML viria com a linguagem XML, uma metalinguagem desenvolvida em 1998 pelo W3C (*World Wide Web Consortium*) com o objetivo de ser uma linguagem com a capacidade de descrever outras *linguagens* de marcação. Com esta característica, que a diferencia da sua irmã HTML, permite-se a organização da informação em um formato estruturado, sem a existência de elementos predefinidos pelo fabricante da linguagem. Essa facilidade abre portas para que usuários dessa linguagem criem sua própria linguagem de descrição documental à cada domínio específico que seja necessário.

Foi baseado na avaliação apresentada acima que elegemos a XML como o instrumento apropriado para o desenvolvimento do nosso trabalho. Utilizando a linguagem XML teremos total liberdade de definir a semântica de nossos marcadores para representação dos registros de programas de disciplinas.

3.1 Como Funciona

A XML é basicamente constituída de um conjunto simples de regras de sintaxe sobre marcação de documentos de texto e dados. São essas marcas que fornecem a estrutura dos dados, portanto é importante compreender que os marcadores criados pelo usuário não fazem parte da linguagem XML. No nosso trabalho criamos alguns marcadores, como por exemplo: **<titulo>**, **<isbn>**, **<autor>**, conforme mostra Figura 1.

```

<? xml version="1.0" encoding="UTF-8" ? >
<UFES>
  <centro> CCJE
    <departamento>Ciência da Informação
      <curso>Arquivologia
        <disciplina>Metodologia Científica
          <programa-disciplina>CIN-38915
            <sugestoes-bibliografica>
              <titulo>Como Elaborar um Projeto de Pesquisa</titulo>
              <sobrenome-autor>Gil</sobrenome-autor>
              <nome-autor>Antonio Carlos</nome-autor>
              <editora>Atlas</editora>
              <edicao>3</edicao>
              <isbn>85-7307-489-2</isbn>
            </sugestoes-bibliografica>
          </programa-disciplina>
        </disciplina>
      </curso>
    </departamento>
  </centro>
</UFES>

```

Figure 1: Metadados que descreve parte da estrutura de nossa base de dados

Um documento XML é estruturado na forma hierárquica, como mostrado no Figura 1 acima. No nosso caso temos uma raiz, que é UFES onde há vários elementos de nível inferior. Subordinados à qualquer um destes elementos podem, por sua vez, conter sucessivamente outros elementos subordinados de níveis ainda mais baixos.

No entanto, como um documento XML contém apenas dados estruturados, para transformarmos o documento em um formato específico em qualquer meio é necessário que utilizemos de algumas tecnologias associadas a linguagem XML. Como no nosso protótipo a intenção foi a de transformar nossa base de dados de programas de

disciplinas em um relatório visível em páginas WEB, utilizamos a linguagem XSLT (*eXtensible Stylesheet Language Transformations*) (Mangano 2002). Esta linguagem tem por objetivo realizar a transformação de dados descritos através de XML em um outro formato. Atualmente é possível transformar-se os dados para os formatos RTF (*Rich Text Format*), PDF (*Portable Document Format*) e até mesmo para um formato em HTML, que é o nosso caso. Além disso, essa linguagem nos permite realização de cálculos com os dados numéricos contidos entre as marcações.

4 Programa de Disciplina: um Meta Documento em XML

Analizando os recursos de representação e processamento de dados disponíveis na tecnologia XML, vemos que os mesmos poderiam ser utilizados para a criação de uma completa base de dados institucional contendo as informações existentes nos programas de disciplinas de uma universidade. Sendo os programas de disciplinas assim armazenados, tais informações poderão ser mais uma boa ferramenta para auxílio no processo de desenvolvimento e avaliação de coleções, como descrito na Seção 2.

Desta forma a proposta neste trabalho foi de um processo de automatização para armazenamento dos programas de disciplinas de cursos de uma IES. Dos programas de disciplinas armazenados poderemos extrair as seguintes informações:

- Indicações das referências bibliográficas dadas pelo corpo docente, os quais são os conhecedores peritos da literatura em suas áreas;
- Necessidades informacionais de diversos cursos oferecidos pela instituição. Estes cursos são frequentados por grande quantidade de usuários da biblioteca;
- O nível de interação entre o corpo discente, docente e a biblioteca, ao se detectar os *caminhos seguidos* por cada uma destas categorias quanto a busca por itens informacionais.

Com estas informações, a biblioteca poderá, com maior propriedade, realizar uma provisão da informação demandada pela comunidade de forma planejada e em tempo hábil.

Para a construção de nosso protótipo de experimentação foram adotadas as seguintes restrições: utilizamos apenas os programas de disciplinas dos cursos de Biblioteconomia e Arquivologia do Departamento de Ciências da Informação da Universidade Federal do Espírito Santo. Destes, coletamos apenas os programas do primeiro semestre letivo de 2004.

Portanto, foram arrolados na amostra, doze programas de disciplinas do curso de Biblioteconomia e apenas um programa de disciplina do curso de Arquivologia, já que muitas das disciplinas do curso de Arquivologia são compartilhadas (ver Tabela 1).

Table 1: Tabela das disciplinas utilizadas no protótipo.

CURSO: Biblioteconomia	
Código	Disciplina
BIB-03888	Normalização da Informação
BIB-03889	Organização e Administração de

BIB-03890	Bibliotecas I Evolução dos Registros do Conhecimento
BIB-03892	Representação Temática I
BIB-03894	Representação Temática II
BIB-03895	Editoração
BIB-03900	Organização e Administração de Bibliotecas II
BIB-03901	Automação de Unidades de Informação
BIB-03902	Métodos e Técnicas de Pesquisa
BIB-03903	Preservação em Unidade de Informação
BIB-03910	Serviço de Recuperação da Informação II
BIB-04994	Pesquisa em Biblioteconomia
CURSO: Arquivologia	
BIB-004431	Mediação e Acesso à Informação Arquivística

Destacamos que essas disciplinas foram utilizadas apenas como modelo para representar o universo de disciplinas oferecidas por uma IES. Em um caso real, teríamos todos os programas de disciplinas de todos os departamentos e cursos na base de dados. Reunindo, desta forma, todas as demandas da instituição em uma base de dados distribuída, porém com processamento podendo ser centralizado, teremos como equacionar a demanda de forma não fragmentada.

Portanto, na representação dos dados precisamos localizar a hierarquia física da instituição em seus nucleares níveis até chegarmos as disciplinas lecionadas pelos cursos. Assim, os dados dos programas de disciplinas representados nesta estrutura, nos possibilitarão inferir informações de demanda a serem confrontadas com o coleção da biblioteca.

No protótipo temos duas estruturas a serem representadas: a primeira contém os programas de disciplinas dos cursos aninhados na estrutura hierárquica-organizacional de uma IES, sendo neste caso representada pela UFES, nos moldes da Figura 1. Uma segunda estrutura, semelhante a primeira, foi criada para representar a biblioteca, compreendendo os itens bibliográficos existentes na coleção.

Observar-se que, ambas estruturas são representadas em uma linguagem XML. No protótipo, colocamos a estrutura com os programas de disciplinas em um arquivo chamado PROGRAMAS.XML e, o coleção da biblioteca, em um outro arquivo de nome ACERVO.XML. Concluída a representação dos dados nos arquivos XMLs, as base de dados serão, ambas, objetos de processamento pela linguagem de transformação XSLT. No experimento, foi utilizada esta última linguagem também para exibir os resultados em uma página WEB. Através deste resultados foi então possível obter-se importantes dados como subsídio aos processos de desenvolvimento de coleções em uma biblioteca universitária, como demonstrado mais adiante.

5 Dando Dinamicidade ao Documento

Como já dissemos na seção anterior, os dados em XML são estruturados de uma forma hierárquica. Cada nível pode ser visto como um nóculo de uma árvore, *desenhada de cabeça para baixo*. O processador da linguagem XSLT nos permitirá *navegar* de um nóculo para outro nesta árvore, obtendo de cada nóculo o dado nele contido. Por exemplo, se quisermos obter o **isbn**, na Figura 1, do livro Como Elaborar um Projeto de Pesquisa, do Antônio Gil, teremos que navegar por todos os **centros**, **departamentos**, **cursos**, **disciplinas** e **programa-disciplinas** para verificarmos se há o livro em alguma das **sugestoes-bibliograficas**.

Em um primeiro momento pode parecer trabalhoso. Porém, não devemos perder de vista que este trabalho será realizado por uma máquina capaz de executar milhões destas atividades em milésimos de segundos.

Todavia, como a navegação por níveis de uma árvore é uma necessidade corriqueira, o W3C definiu facilidades na linguagem XSLT capazes de prover formas de navegarmos por entre nóculos de um mesmo nível hierárquico sem que, para isso, precisemos acessar os níveis superiores da árvore cada vez acessarmos um nóculo inferior. Em outras palavras, através dessa facilidade seremos capazes de navegarmos por todas as **sugestoes-bibliograficas** de uma só vez em busca do item, ou itens, informacional desejado.

A XSLT é portanto uma boa ferramenta de transformação de documentos estáticos em um outro documento. Dando assim um caráter dinâmico da informação. Neste trabalho estamos interessados em transformar os documentos estáticos dos programas de disciplinas em informação dinâmica para ser divulgado aos gestores da coleção. Essa dinamicidade nos permite diversas contextualizações, ou modos diferentes de ver um mesmo documento se desdobrar em vários outros. Dessa forma pretendemos dar uma visão geral de como foi utilizada essa linguagem, a XSLT, para a realização dessas diversas contextualizações aplicadas nesse trabalho. Não entraremos em detalhes da linguagem propriamente dita, para que o texto não fique enfadonho. Contudo, recomendamos para o leitor mais interessado, algumas referências especializadas neste assunto (Ray 2003, Mangano 2002), ou mesmo o trabalho (Santos 2005) para maiores detalhes de implementação.

Quando processados os conteúdos dos programas de disciplinas existentes na base de dados do protótipo e contraposta com a base de dados que representa o acervo de nossa biblioteca, o que temos é a transformação de nossos dados em informação. Portanto, a partir destas estruturas, usamos um detalhamento das sugestões bibliográficas contidas nos programas de disciplinas do curso de Biblioteconomia, a serem representadas nas seguintes categorias.

- *Caracterização da demanda*: Década de publicação; Formato do material; Identificar o tipo de bibliografia (Básica e Complementar) recomendada nos programas de disciplinas;
- Comparação da demanda bibliográfica com o acervo de uma biblioteca universitária:
 - A relação entre disciplina e acervo; A classificação por tipo de bibliografia (Básica e Complementar) das disciplinas e comparação com o acervo;
 - Avaliação do grau de atendimento por década.

Identificamos, no universo representado pelos 12 programas de disciplinas do curso de Biblioteconomia da UFES, um conjunto de 225 sugestões para diversos formatos bibliográficos (Livros, periódicos, monografias, etc.). Sendo que desse total observa-se

um grande percentual de sugestão para a disciplina Organização e Administração de Bibliotecas II (OABII), com 28,44% do total, com 64 sugestões de itens bibliográficos, veja Tabela 2.

Table 2: Quantidade de itens sugeridos por disciplinas.

CURSO: Biblioteconomia			
Código	Disciplina	Sugestões	%
BIB-03901	Automação de Unidades de Informação	3	1.33
BIB-03895	Editoração	17	7,56
BIB-03890	Evolução dos Registros do Conhecimento	21	9.33
BIB-03902	Métodos e Técnicas de Pesquisa	5	2.22
BIB-03888	Normalização da Informação	11	4.89
BIB-03889	Organização e Administração de Bibliotecas I	19	8.44
BIB-03900	Organização e Administração de Bibliotecas II	64	28.44
BIB-03903	Preservação em Unidade de Informação	33	14.67
BIB-03892	Representação Temática I	24	10.67
BIB-03894	Representação Temática II	20	8.89
BIB-03910	Serviço de Recuperação da Informação II	1	0.44
Total Geral		225	100.00

O resultado acima explicita as indicações feitas por cada disciplinas que escolhemos para fazer parte de nossa base de dados protótipo. Entretanto, se olharmos para o curso de Biblioteconomia como um todo, naquele semestre, veremos que o total de itens

sugeridos difere do total de 225 acima apresentado. Isso deve-se ao fato de que, no resultado acima, não olhamos a duplicidade de indicação de itens informacionais. Uma disciplina pode indicar um item também indicado por outra disciplina.

Notamos, em nossas análises, uma predominância de 49,26% para o formato livro. Foram 100 sugestões em relação ao total geral de 203 sugestões. As distribuições para os outros suportes ficaram da seguinte forma. Para periódico impresso ocorreram 52 sugestões, o que representa 25,62% do total das sugestões. Em contra-partida, para periódico eletrônico forma 1,48%, com 3 sugestões do total das sugestões. Documento eletrônico representou 7,88%, sendo 16 sugestões do total. Anais com 21 sugestões do total das sugestões representando 10,34%; Fita VHS representou apenas 2,96%, com 6 sugestões. CD-ROM representou 0,49% e monografias foram indicadas apenas 4, representando 1,97% total geral.

Destas constatações estatísticas observamos que o formato livro com 49,26%, ainda é o principal elemento bibliográfico preferido pelo corpo docente representando quase metade das sugestões elaboradas pelas 12 disciplinas do curso de Biblioteconomia da UFES.

Fizemos também uma caracterização por ano de publicação dos itens sugeridos. Neste exercício, as publicações foram sistematizadas pelas categorias: década (70, 80) e quinquênio 90-94, 95-99 e 00-04. A análise das sugestões bibliográficas apresentou uma predominância de 32%, ou seja, 74 de um total de 100 sugestões, para publicações feitas a partir do ano de 1995. Este índice tem um peso mais significativo para o quinquênio 2000-2004, com 39 indicações. Isto constata que o corpo docente tem procurado acompanhar as publicações mais recentes da área. As sugestões bibliográficas relativas ao quinquênio 1990-1994, respondem pelo percentual de 14%. Isto representa um pouco menos que a metade das sugestões bibliográficas que compreendem o período de 1995 em diante. Notamos que a soma das sugestões bibliográficas compreendidas entre as décadas de 70 e 80 totalizam juntas menos que o percentual recomendado para publicações do período 1990-1994. No entanto ressaltamos que estas são indispensáveis para compor a bibliografia básica do curso de Biblioteconomia.

Muitas outras análises podemos fazer dos dados contidos nos programas de disciplinas. Um estudo mais aprofundado, inclusive com descrição dos códigos que usamos para gerar esta e outras análises podem ser encontradas em (Santos 2005).

6 Conclusões

Apresentamos neste trabalho uma forma adicional de obtenção de dados quantitativos da necessidade informacional do usuário de uma biblioteca universitária. Tendo em vista o número reduzido de pessoal, realidade enfrentada por muitas bibliotecas hoje, nossa proposta também aponta uma estratégia de automação desse procedimento.

Utilizamos a linguagem XML para a construção de uma base de dados de programas de disciplinas, bem como para simular a coleção de uma biblioteca real. Através disso, utilizamos a linguagem de transformação de dados XSLT para o processamento e geração automática de páginas WEB com diversas análises de demandas de itens para a coleção de nossa hipotética biblioteca.

Entendemos que a solução por nós aqui apresentada, quando agregada à outras já apresentadas na literatura, auxiliará grandemente na identificação da demanda informacional para uma gestão mais fecunda da coleção.

Um dos próximos objetivos de continuidade desse trabalho é a integração, via tecnologia XML, com livrarias *on-lines*, tal como a **amazon.com**. Pois através dessa integração poderemos instantaneamente saber quanto custa atender a demanda de atualização para a biblioteca. Mais que isso, saber quanto custa uma obra ou mesmo todo o acervo, em um dado momento.

Referências

- DANILOWICZ, C.; SZARSKI, H. Selection of Scientific Journals Based on the Data Obtained from an Information Service System. *Information Processing & Management*, v. 17, n. 1, p. 13–19, 1981. Disponível em: <<http://www.sciencedirect.com/science/article/B6VC8-469KM57-3/2/af314a7434ef5ae935ee0fe84451654f>>.
- KANTARDZIC, M. *Data Mining : Concepts, Models, Methods, and Algorithms*. New York: IEEE, 2002.
- KAO, S. C.; CHANG, H. C.; LIN, C. H. Decision Support for the Academic Library Acquisition Budget Allocation via Circulation Database Mining. *Information Processing & Management*, v. 39, n. 1, p. 133–147, 2003. Disponível em: <<http://www.sciencedirect.com/science/article/B6VC8-452F87Y-1/2/ea486004005f96a70256b16902db6a97>>.
- MANGANO, S. *XSLT Cookbook*. 1. ed. New York: O'Reilly, 2002.
- OLIVEIRA, E. Identificando Necessidades de Acervo de Usuários de um Departamento Acadêmico: uma Abordagem Bibliométrica. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 126–131, 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n1/v33n1a15.pdf>>.
- RAY, E. T. *Learning XML*. 2. ed. New York: O'Reilly, 2003.
- SANTOS, M. H. dos. *Desenvolvimento de Coleções: Uma Proposta Automatizada à Avaliação de Coleções em Bibliotecas Universitárias*. Vitória, ES, 6 2005.