

HEMEROTECA DIGITAL DO CMU: PROJETO PILOTO

Autor:

¹Elizabeth Maria Alcântara P. Pazini – emapp@unicamp.br

²Rosaelena Scarpeline – rscarpel@unicamp.br

*³

Resumo:

Relato da experiência da Biblioteca do CMU, na elaboração, montagem, indexação de uma hemeroteca, estabelecendo metodologias para que a partir de uma análise criteriosa dos artigos, seja desenvolvido um conjunto de padrões para identificar, armazenar e disponibilizar as informações em texto completo dos artigos de jornais e revistas que compõe seu acervo num ambiente digital, possibilitando a busca e a recuperação deste tipo de material não só através de índices, thesaurus e base de dados, mas também a pesquisa palavra por palavra, resultando na interoperabilidade entre interfaces abertas e protocolos de comunicação de dados, onde os usuários poderão realizar suas consultas de forma unificada em uma única interface, sem interferência de terceiros, fornecendo também a possibilidade de retornar o texto digital em papel.

Palavras-chave: Biblioteca digital; Hemeroteca

¹ Bibliotecária responsável pelo setor de Periódico e Hemeroteca

² Bibliotecária Diretora da Biblioteca

³ Biblioteca do Centro de Memória da Unicamp. Caixa Postal 6123. Campinas – S.P. 13.083-970 - Brasil

INTRODUÇÃO

A Biblioteca do Centro de Memória possui uma Hemeroteca com cerca de 45.000 recortes de jornais e revistas diversos, com interesse especial para a história e o cotidiano da cidade de Campinas. Abrange um período que vai de 1920 até os dias atuais, divididas em duas grandes coleções:

Coleção João Falchi Trinca: o maior bibliófilo da história de Campinas. Composta de 18.000 artigos de jornais e revistas de Campinas e região, retratando o cotidiano da cidade durante mais de 70 anos.

Coleção Corrente: com o passar dos anos fomos recebendo mais coleções que foram compondo essa Hemeroteca. Também passamos, desde 1986, a colecionar os jornais diários da cidade, selecionando as matérias de interesse e anexando ao acervo.

Com o apoio da FAEP e da FAPESP a coleção João Falchi Trinca foi trabalhada física e tecnicamente, e como resultado deste trabalho inicial, publicamos em 1997 o índice/tesaurus “História de Campinas através da Hemeroteca João Falchi Trinca: descritores e afins”, na série Instrumentos de Pesquisas do CMU, trazendo os assuntos indexados de forma ordenada, respeitando a linguagem natural uma relação dos principais assuntos tratados nesta coleção, facilitando assim o processo de consulta.

Além de classificar os artigos pelos seus assuntos específicos, este sistema foi elaborado para facilitar as pesquisas no acervo da hemeroteca.

A séries são as seguintes:

1. Aspectos urbanos;
2. Ciências;
3. Cultura
4. Economia;
5. Educação;
6. História
7. Literatura;

8. Personagem;

9. Política;

10. Religião;

11. Transporte

Cada série subdivide-se em sub-séries, classes, sub-classes, grupo, sub-grupo etc, quantas vezes se fizer necessário.

Com o Know-how adquirido neste trabalho, passamos a trabalhar então a outra coleção e partindo do índice já implantado fomos desenvolvendo os temas específicos e criando novas palavras chaves.

Diante do grande volume de artigos e de informações geradas através dos índices surgiu a necessidade de se estudar uma maneira fácil e rápida de disseminação dessas informações.

Em 1998, novamente com o auxílio da Fapesp desenvolvemos as bases de dados on-line, em Microisis, que são roteadas para internet pelo software WWWIsis, visando uma rápida recuperação de dados, permitindo uma pesquisa mais específica, seja por autor, título, assunto específico. Atualmente estamos on-line com 38.000 artigos.

DIGITALIZAÇÃO

Sendo a internet um espaço muito utilizado na disseminação da informação, tanto para disponibilizar os serviços tradicionais como para a inclusão das novas formas de busca, há cada vez mais por parte dos pesquisadores anseios quanto ao uso dos serviços eletrônicos.

O serviço de referência hoje, porta de entrada para os acervos das bibliotecas passaram a funcionar no ambiente virtual com uma precisão absoluta. Desta forma, fomos impelidos a procurar novas ferramentas que dessem ao pesquisador uma maior abrangência no momento da busca, não só através de autoria e assunto específico, mas a liberação do texto integral,

possibilitando a varredura do texto, palavra por palavra, sem que ele precise da interferência de terceiros.

A partir da inauguração da Biblioteca Digital da Unicamp, inaugurada em maio de 2002, gerenciada pelo Sistema Nou-Rau, que implementa um serviço on-line de armazenamento e obtenção de documentos digitais, promovendo um acesso controlado e mecanismos eficientes de busca, passamos a estabelecer os primeiros contatos para o estudo de parceria onde o objetivo seria a colocação do texto integral dos artigos que compõem o acervo da Hemeroteca. O Sistema Rau-Tu (Nou-Rau) foi desenvolvido pelo Instituto Vale do Futuro em parceria com o Centro de Computação da Unicamp. É de uso livre e gratuito.

Começamos então por avaliar a transformação da linguagem da base atual Microsis para ASCII com a finalidade de obter suporte com o protocolo Z39.50. Este protocolo era a exigência do Sistema Nou-Rau, transformando a base de Microsis numa linguagem compatível com o próprio sistema, não havendo necessidade da conversão da base de dados como um todo.

OBJETIVOS

Tendo como noção básica que o usuário tem que estar sempre bem orientado, sendo capaz de navegar pelas bases de informação com segurança procuramos montar um site de pesquisa levando em conta:

- Disponibilização da informação durante 24hs. permitindo acesso em sua própria residência ou em lugares onde não poderá obter orientação próxima;
- Agilidade para a obtenção da informação;
- Uso simultâneo do mesmo documento por vários pesquisadores
- Preservação dos originais

METODOLOGIA

Tomou-se por base a especificidade da biblioteca para podermos ter um ponto de partida. Escolhemos para esta primeira etapa, como projeto piloto, a *Série História*, pois traz em sua formação temática varias fontes que norteiam a pesquisa em história de Campinas, já que o perfil de nossa biblioteca é história regional.

A parceria CMU/BC para a Biblioteca Digital começou com o levantamento de procedimentos necessários para que o trabalho pudesse dar seguimento, neste momento contamos não só com o apoio na disponibilização de software compatíveis, mas também com o empenho do Diretor de Sistema Automatizados Marcos Dario G. Sae e dos estagiários de informática Felipe Rodrigues Uriane e Marcelo S. Macedo.

Partindo da biblioteca on-line hemeroteca, e com os artigos trabalhados física e tecnicamente iniciamos os passos seguintes:

- 1 – Transformar a linguagem microsis para ASCII
- 2 – Digitalizar os textos em PDF
- 3 – Digitalizar as imagens em JGJ – 150 DPIs
- 4 – Anexar texto à imagem
- 5 – Inserção do arquivo na base Rau-Tu
- 6 – Reindexação da base Rau-Tu feita a cada 2 horas

No momento da inserção de cada texto digitalizado em PDF na base Rau-Tu, segue o MFN correspondente, desenvolvido em microsis que possibilita a pesquisa também pelo registro bibliográfico na Biblioteca Digital.

A base de dados da Hemeroteca na Biblioteca on-line CMU em Wisis permanece com sua integridade, apenas será acrescentado um link no MFN digitalizado, informando que o referido artigo está disponível em texto integral na Hemeroteca Digital.

Ao se clicar neste link o usuário terá acesso de imediato ao texto em PDF, não havendo necessidade de reformular sua pesquisa na Biblioteca Digital, obtendo assim uma resposta mais ágil e eficaz para sua pesquisa.

Desta forma fica disponível na rede Web duas formas de pesquisa:

1. Dentro da Biblioteca on-line CMU, na base de dados Hemeroteca por autor, título e assunto específico;
2. Dentro da Biblioteca Digital da Unicamp, na Hemeroteca Digital, onde além da pesquisa disponível na biblioteca on-line, poderá também fazer a varredura no texto integral, palavra por palavra.

As bases de dados, Biblioteca on-line Hemeroteca e Hemeroteca Digital estão abrigadas em servidores diferentes, no primeiro caso em um servidor Pentium III 500 na própria Biblioteca do CMU, e no segundo um servidor Dual Pentium III 750 mhz do Sistema de Biblioteca da Unicamp, no Centro de Computação (CCUEC).

Durante o planejamento dos trabalhos alguns pontos foram relevantes para a tomada de decisões em relação ao software utilizado

- Capacidade de gerenciar grandes massas de dados;
- O grande volume de informação não afetaria a velocidade nem a performance do sistema, no que se refere a recuperação dos dados;
- A pesquisa é realizada palavra por palavra buscando informação no texto integral e também por assunto específico com expressões booleanas;
- Facilidade para acessar as bases sem a necessidade de conhecimentos específicos em informática.

Outros pontos que foram levados em consideração foi a instalação de uma interface amigável que permitisse a facilidade do manuseio e a que fosse garantido a proteção das bases contra usos indevidos do sistema.

Para se obter uma resposta rápida no momento de se visualizar o texto integral é necessário uma arquitetura mínima:

1. Computador Pentium,
2. Monitor colorido SVGA 800X600 256 cores;
3. Impressora com resolução igual ou superior a 150 dpi;
4. Sistema operacional Windows 95/98 ou outras versões atualizadas do mesmo;
5. Internet Explorer versão 5 ou similar;
6. Software Adobe Acrobat Reader.

Conclusão

Com esse trabalho esperamos colocar a disposição dos pesquisadores facilidades que ultrapassem o modo tradicional de pesquisa em papel, facilitando sua interação remota entre a biblioteca e seus usuários.

A disponibilização do texto integral no ambiente eletrônico representa uma economia de tempo e esforço no momento da realização de uma pesquisa, não só para o pesquisador mas também para o bibliotecário. Isso só é possível se o usuário estiver bem orientado para saber com segurança levantar palavras chaves sem necessitar de um atendimento mais refinado e direto, para que a resposta obtida seja rápida, limpa e devidamente filtrada.

Para isso estamos trabalhando no desenvolvimento de um browser para o recebimento de mensagens, procurando medir a satisfação do usuário, receber sugestões ou fazer uma orientação mais aprofundada na pesquisa a ser desenvolvida, esse será nosso serviço de referencia on-line dentro da base específica de hemeroteca, interagindo diretamente bibliotecário de referencia X usuário on-line. Desta forma as dificuldades serão detectadas e o serviço oferecido poderá ser melhorado.

Não podemos deixar de lembrar que a disponibilização da Hemeroteca Digital estará funcionando também como um serviço de marketing para os

outros serviços da biblioteca do Centro de Memória e suas bases de dados online, sendo o ponto de partida para que possamos estudar a possibilidade de transformarmos outras bases em bibliotecas híbridas e interativas.

Para o acesso da Biblioteca online do CMU Hemeroteca
<http://www.unicamp.br/suarq/cmu/biblioteca>

Para o acesso da Biblioteca Digital da Unicamp
<http://www.unicamp.br/bc/bibdig/index.htm>

REFERÊNCIAS BIBLIOGRÁFICAS

ASHWOTH, W. **Manual de bibliotecas especializadas e serviços informativos**. 2.ed. Lisboa: Fund. Calouste Gulbenkian, 1981.

BELLOTTTO, H. L. Arquivos, bibliotecas e museus: fronteiras definidas. **Arquivo Rio Claro**, Rio Claro, v.1, n.1, p.5-11, jan. 1982.

CAVALCANTI, C. R. **Indexação & tesauro**: metodologia & técnicas. Brasília: ABDF, 1978.

CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. São Paulo: APB, 1994.

CÓDIGO de Catalogação Anglo-Americano. 2.ed. São Paulo: FEBAB, 1983-1985.

COELHO, G. M. et al. **Elaboração de um “thesaurus” experimental para o setor siderúrgico**. Rio de Janeiro: IBS, 1976.

COULON, D.; Kayser, D. **Informática e linguagem natural**: uma visão geral dos métodos de interpretação de textos escritos. Brasília: CNPq; IBICT, 1992.

FERREZ, H. D.; BIANCHINI, M. H. S. **Thesaurus para acervos musicológicos**. Rio de Janeiro: Fund. Nacional Pró-Memória, Coord. Geral de Acervos Musicológicos, 1987.

FIGUEIREDO, N. M. **Serviço de referência e informação**. São Paulo: Polis: APB, 1992.

FUJITA, M. S. L. **Elaboração de tesouros monolíngues com o programa tecer**: considerações sobre o uso. São Paulo: APB, 1997.

GOMES, H.; GUSMÃO, H. **Guia prático para elaboração de índice**. Niterói: GBICDSH, 1983.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Diretrizes para elaboração de thesaurus monolíngues. Brasília: IBICT, 1984.

INSTITUTO CULTURAL ITAÚ. Banco de dados informatizados: módulo fotografia, setor memória fotográfica da cidade de São Paulo. São Paulo: ICI, 1992.

LANCASTER, F. W. **Construção e uso de tesouros: curso condensado**. Rio de Janeiro: IBICT, /1980?/.

_____. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos/Livros, 1993.

MARCHIORI, Patrícia Z. "Ciberteca" ou Biblioteca Virtual: uma perspectiva para o gerenciamento de recursos de informação. **Ciência da informação**, v. 26, n. 2, 1997. p.115-27.

MARCONDES, Carlos Henrique, GOMES, Sandra Lúcia. O impacto da internet nas bibliotecas brasileiras. **Transinformação**, v.9, n.2, 1997. (on-line)

_____, SAYÃO, L.F. **Projeto técnico da biblioteca digital brasileira em C&T**. (online). Disponível em <http://www.ibict.br> Acesso em 01 abr. 2002.

MARTINS, M. G. Recortes: esses precisos desprezados. **Revista do Arquivo Público**, Recife, 1981-2.

PAES, M. L. **Arquivo: teoria e prática**. Rio de Janeiro: FGV, 1986.

PAZIN, R. A. A. P. **Indexação de multimeios**. Curitiba: Scientia et labor, 1988.

ROBREDO, J. **Documentação de hoje e de amanhã**: uma abordagem informatizada da biblioteconomia e dos sistemas de informação. 2.ed. Brasília: Autor, 1986.

SÁ, E. S. (Coord.) et al. **Manual de normalização e trabalhos técnicos, científicos e culturais**. Petrópolis: Vozes, 1994.

SCARPELINE, R. Hemeroteca Centro de Memória – Unicamp. Simpósio de Técnicos de Apoio ao Ensino e à Pesquisa, **Anais ...1**, 1997, Campinas. Caderno de resumos.

_____, coord. **História de Campinas através da hemeroteca João Falchi Trinca**: descritores e afins. Campinas: Área de Publ. – CMU, 1997.

SCHELLENBERG, T.R. **Arquivos modernos**: princípios e técnicas. Rio de Janeiro: FGV, 1974.

SILVA, E.P. **Serviço online de referência e informação**. São Paulo: ECA/USP, s.d. Disponível em: <http://www.referenciaonline.hpg.ig.com.br/index.htm>. Acesso em 11 abr. 2002

VIET, J. **Thesaurus internacional do desenvolvimento cultural**. Rio de Janeiro: Fundação Casa de Rui Barbosa, 1983.

ZAMBEL, M. M. **Glossário de termos usuais em biblioteconomia e documentação**. São Carlos: Fund. Theodoretto Souto, 1978.

ZANAGA, M. P. **Referenciação bibliográfica**: aplicação da NBR 6023/89. Campinas: PUCCAMP/FABI, 1992.