

Automatização de documentação não convencional um estudo de caso em engenharia

Brasilina Passarelli
Doutoranda – ECA/USP

Resumo – O projeto para “Editoração e Recuperação Automatizadas de Curricula-Vitarum” foi desenvolvido para uma empresa brasileira de engenharia consultiva a fim de otimizar o tempo gasto para seleção de equipes técnicas visando a participação da mesma em licitações. Foi concebido prevendo a criação de dois sistemas: um para a editoração automatizada dos textos dos curricula, utilizando um software editor de textos, e um segundo sistema estruturado com a filosofia de bancos de dados, para permitir a realização de buscas a partir de parâmetros isolados ou em conjunto.¹

1 Introdução

O projeto para “Editoração e Recuperação Automatizada de Curricula-Vitarum” foi originalmente desenvolvido para uma empresa brasileira de engenharia consultiva, com o objetivo de possibilitar à Diretoria Comercial da mesma selecionar, rápida e eficazmente, profissionais universitários integrantes do seu quadro de funcionários efetivos para composição de equipes técnicas, visando a participação da empresa em licitações, além de servir como instrumento auxiliar na tomada de decisões referentes aos processos de movimentação interna e recrutamento externo de pessoal.

As áreas de atuação da consultora abrangem estudos, planejamento e projetos em todos os campos da engenharia civil, ambiental, elétrica e mecânica, da arquitetura, da economia e do gerenciamento global de empreendimentos. Dispõe de aproximadamente 3.500 funcionários, dos quais cerca de 1.100 são de nível universitário.

A fim de dar suporte às atividades comerciais da empresa, definidas e gerenciadas pela sua Diretoria Comercial, foi criado nesta Diretoria um Centro de Documentação, responsável pelas coletas, produção, armazenamento, recuperação e disseminação das informações técnico-comerciais da empresa, necessárias à elaboração de cadastros e preparo de documentação para habilitação em licitações junto a Órgãos Governamentais.

¹ Projeto apresentado como DISSERTAÇÃO DE MESTRADO à ECA/USP. sob orientação do Prof Dr Frederic Michael Litto em outubro de 1987

AUTOMATIZAÇÃO DE DOCUMENTAÇÃO NÃO CONVENCIONAL

Tanto para CADASTROS como para HABILITAÇÃO EM LICITAÇÕES é exigida das empresas a apresentação de documentação relativas às: capacidade jurídica; capacidade técnica, idoneidade financeira e regularidade fiscal. O “dossiê” para comprovação da capacidade técnica, normalmente o mais volumoso, é tradicionalmente composto de: Curricula-Vitarum dos integrantes da Equipe Técnica Permanente da Empresa; Histórico dos Serviços já Executados e em Execução pela Empresa, além da apresentação de Atestados de Capacidade Técnica referentes aos serviços desenvolvidos, os quais são emitidos pelos respectivos clientes.

2 Justificativa

Considerando-se que dos aproximadamente 1.100 curricula cada um possuía em média 10 fls. tamanho A-4; que seu conteúdo podia ser alterado em função da comprovação de um determinado tipo de experiência requerida nos editais e, quando isto ocorria, quase sempre era necessário que o mesmo fosse totalmente redatilografado; que eles eram exaustivamente acessados para obtenção de cópias xerox e este procedimento acabava deteriorando o papel vegetal utilizado como suporte para datilografia e, finalmente, que a cada atualização anual de todos os curricula, o volume de datilografia, conferências e correções requeria a alocação de 03 (três) datilógrafas a 160h/mês por 06 (seis) meses, iniciamos o desenvolvimento do projeto para automatização da editoração e recuperação dos curricula.

3 Metodologia

Uma revisão de literatura abrangendo o período 1975-1985 realizada através do DIALOG Information Retrieval Service, acessando as bases de dados do LISA (Library and Information Science Abstracts), não nos trouxe notícias acerca de sistemas automatizados de curricula.

Outras empresas de engenharia consultiva foram também contatadas, em São Paulo, com o intuito de se detectar experiências similares em curso. Entretanto esta investida também não trouxe resultados, devido ao sigilo com que estes assuntos são tratados nas empresas.

O projeto para automatização dos currículos foi desenvolvido prevendo-se a criação e implantação de dois sistemas em duas etapas distintas.

3.1 Modelo Proposto I - Sistema para Editoração Automatizada dos Curricula

A configuração adotada para o desenvolvimento e implantação do sistema de editoração automatizada de curricula compreende um computador

BRASILINA PASSARELLI

IBM 4341 com impressora, operado em regime de multiprocessamento, a alocação de um disco IBM 3375 com capacidade de memória de 05 cilindros e um terminal de vídeo IBM 3278.4.

O software utilizado foi o SCRIPT/VS da IBM, um processador de palavras, gerenciado pelo CMS/IBM - Conversational Monitoring System.

Para otimização dos recursos oferecidos pelo SCRIPT/VS, realizou-se uma revisão no padrão de apresentação dos currícula gerados manualmente, a nível de conteúdo informacional e formatação estética dos mesmos, com o intuito de se definir o padrão para os currícula editorados via computador.

Foram definidos como campos obrigatórios: Nome do Funcionário; Data e Local de Nascimento; Instrução (formação superior básica); Línguas; Associações . Profissionais (é obrigatório o n- do CREA – Conselho Regional de Engenharia, Arquitetura e Agronomia, uma vez que 80% do corpo técnico da empresa é constituído por engenheiros); Cargo Atual na Empresa (descrito em ordem hierárquica crescente); Histórico Profissional: este campo, normalmente o mais longo, foi limitado de acordo com o tempo de formado de cada profissional, em função de se reduzir o custo da locação do espaço em disco necessário para o sistema, de acordo com os seguintes parâmetros: formado até 05 anos - até 03 folhas A4; formado de 05 a 10 anos – até 06 folhas A4; formado há mais de 10 anos – até 08 folhas A4.

Foram considerados como campos não-obrigatórios: Cursos Especiais e/ou de Pós-Graduação; Publicações e Atividades Docentes. O campo Congressos e Seminários foi suprimido considerando-se que aqueles que tiveram uma participação efetiva nos mesmos apresentaram trabalhos. Assim sendo estes serão listados no campo Publicações.

O Sistema de Editoração Automatizada de Currícula apresenta em seu MENU, as seguintes opções: 01 – Pesquisa no Cadastro de Currícula; 02 – Inclusão no Cadastro de Currícula; 03 - Alteração de Dados do Cadastro; 04 – Impressão do Cadastro de Currícula; 05 – Recuperação de Currícula da Fita; 06 - Recuperação de Currícula da Fita para impressão; 07 - Leitura da fita; 08 - Leitura da Fita para impressão; 09 - Inclusão de Currícula novos na Fita; 10 - Atualização da Fita; 11 - Controle do Sistema; 12 - Cadastramento de Usuários no sistema; 13 – Recuperação de Diretório de Fita; 14 – Restauração de Diretório de Fita; 15 – Impressão de Currícula em Disco.

A opção 01 permite a pesquisa pelo nome do funcionário combinada em até 03 formas. A primeira pelo nome completo do funcionário; a segunda pelo primeiro nome e pelo menos um sobrenome, para identificação de

AUTOMATIZAÇÃO DE DOCUMENTAÇÃO NÃO CONVENCIONAL

uma mesma pessoa, cujo nome completo é desconhecido, usando-se o conector AND e, a terceira, pelos pré-nomes de duas pessoas distintas, usando-se o conector OR. As buscas nesta pesquisa obedecem a parâmetros desenvolvidos utilizando-se lógica booleana. A pesquisa pelo nome do funcionário é a única, neste sistema, desenvolvida para atendimento de usuários. As outras opções foram criadas para facilitar as rotinas operacionais e de gerenciamento do sistema. A Figura 1 apresenta um Fluxograma Geral do Sistema de Editoração dos Currícula.

3.2 Modelo Proposto II – Recuperação por Buscas Combinadas

Uma vez implantado o sistema para editoração automatizada dos textos dos currícula, fez-se necessário o desenvolvimento de um segundo sistema, capaz de permitir a realização de buscas por parâmetros isolados ou em conjunto, para responder a perguntas do tipo: “preciso de especialistas em mecânica dos solos, com mais de 10 anos de formados, experiência mínima em cargo de chefia e/ou coordenação de projetos de 3 anos”. Desta forma, o sistema deveria estar apto a prover resumos dos currícula selecionados na busca e, então, de posse dos mesmos, o usuário solicitante avaliaria os resultados obtidos e requisitaria cópias impressas do texto completo dos currícula selecionados.

Baseados nas informações contidas nos textos do sistema de editoração de currícula, foram definidos formulários de entrada de dados. O software escolhido para gerenciamento dos dados, dentre aqueles disponíveis no mercado e na empresa e compatível com a configuração adotada (definida no MODELO PROPOSTO I) foi o STAIRS/VS-IBM (Storage and Information Retrieval System/Virtual Storage). A figura 2 apresenta a Arquitetura do Sistema de Recuperação por Buscas Combinadas.

O STAIRS/VS trabalha com conceitos de campos não-formatados (parágrafos) e campos formatados (fixos).

Foram definidos como campos não-formatados, aqueles de informação textual podendo variar até um limite estabelecido, os seguintes: Instrução; Tipos de Atividades; Áreas de Atuação; Pós-Graduação; Línguas; Clientes para os quais já trabalhou (fora da empresa); Clientes para os quais já trabalhou (dentro da empresa).

Como campos formatados, textuais ou numéricos, foram considerados: Nome do Funcionário; Código da Área (informação localizacional); Tempo de Formado (ano de formatura); Chefia (fora da empresa - somatória dos anos de experiência); Chefia (dentro da empresa - a partir do primeiro ano de experiência); Extensão Universitária (sim ou não); Docência (sim ou não); Publicações (quantidade).

A nível de preparo das informações para entrada no sistema, a maior dificuldade surgiu em função da terminologia a ser utilizada para indexação do campo *Áreas de Atuação*. Com o intuito de elaborar uma lista de descritores que refletisse a qualidade técnica da empresa e de seu acervo humano (seu corpo técnico) efetuou-se um levantamento entre as superintendências técnicas, solicitando que cada uma indicasse até 30 (trinta) descritores relativos às suas áreas de atuação. As respostas foram então compatibilizadas, com vistas principalmente à supressão de duplicidades e estabelecimento de termos sinônimos. Como resultado estruturou-se um arquivo do tipo dicionário, tendo como finalidade principal a manutenção da uniformidade das entradas. A hierarquização dos descritores não foi efetuada, uma vez que o STAIRS/VS opera buscas pelo radical, sendo que as sentenças de buscas são equacionadas utilizando-se linguagem natural aliadas aos conectores. Dentre as 8 funções apresentadas pelo software, 3 delas são especificamente de buscas e estão baseadas em diferentes métodos: SEARCH - pesquisa as informações apresentadas nos campos não-formatados; SELECT - pesquisa as informações apresentadas nos campos formatados; RANK - classifica as buscas em até 05 níveis de hierarquia.

Vamos agora, a título de exemplo, elaborar a sentença para a busca definida na introdução do MODELO PROPOSTO II:

Palavras-chave: Microcomputador; Ensino de Biblioteconomia

“Preciso definir uma equipe composta de especialistas em Mecânica dos Solos, com mais de 10 anos de formado e experiência mínima em cargo de chefia e/ou coordenação de projetos de 3 anos”.

1. ..search mecanica with solos. assuntos.
2. ..select 1 formado > 10.
3. ..select 2 chefia >= 3.

Com 3 (três) sentenças teremos acesso a todos os currícula do banco de dados que atendam aos parâmetros especificados. As solicitações de buscas podem ser altamente específicas ou mais genéricas, de acordo com as especificações dos editais de licitação ou de cadastramento de empresas para posterior prestação de serviços. É importante ressaltar que quanto melhor definidos os parâmetros de busca, mais relevantes serão as respostas obtidas.

4 Análise Comparativa entre as “Performances” dos Sistemas Manual X Automatizado

4.1 Editoração

Quando manualmente executados os currícula eram datilografados em vegetal e arquivados em pastas suspensas verticais, por ordem alfabética de so-

AUTOMATIZAÇÃO DE DOCUMENTAÇÃO NÃO CONVENCIONAL

brônimos. A cada atualização anual eram inteiramente redatilografados. Os vegetais de tanto serem manuseados, para obtenção de cópias xerox, acabavam por se deteriorar.

O sistema de editoração automatizada exige, para digitação e operação um treinamento especial. Em contrapartida, as atualizações são rápidas uma vez que a exclusão, movimentação de linhas ou parágrafos e inclusão são feitas através de comandos especiais. O arquivamento, originalmente em pastas suspensas foi substituído por fitas e/ou disco magnético. Afora a economia de espaço físico, superou-se também o problema de deterioração dos vegetais, uma vez que os suportes magnéticos atualmente utilizados são seguramente mais duráveis.

4.2 Recuperação

No sistema manual, a única possibilidade de busca possível era pelo nome completo, já que nenhum tratamento de indexação era dispensado aos currícula. A seleção de elementos para composição de equipes técnicas somente podia ser feita pelos dirigentes das áreas técnicas, uma vez que eram os que melhor conheciam os profissionais a eles afetos.

O sistema de recuperação automatizada permite uma grande variação de combinações de busca e proporciona ao usuário o acesso, tanto “on-line” como através de relatórios impressos, a uma espécie de “abstract” dos textos dos currícula. Desta forma, o sistema funciona para a Diretoria Comercial como um instrumento único para tomada de decisão quando da seleção de equipes técnicas, instrumento este totalmente inexistente no sistema manual que o precedeu.

Abstract – A project entitled “Automatic Editing and Retrieval of Curricula-Vitarum” was developed for a large Brazilian engineering consulting company, with the purpose of optimizing the process of proposal preparation. Automatization was developed in two steps. The first providing automatic editing of curricula texts, utilizing a word processor, and the second, structured as databank to provide searches within isolated parameters as well as composed ones.

Palavras-Chave (Key Words): Automatização (Automatization), Documentação não Convencional (Non Conventional Documentation), Curricula-Vitarum, Editores de Textos (Word Processors), Recuperação Automatizada (Automatic Retrieval) e Gerenciadores de Banco de Dados (Data Bank Management Systems). (BP)

5 Bibliografia

1. BOEHME, F. & WERNER, E. Computer-aided search method and machine-readable data banks developed and used by IBM Germany Patent Documentation. *WORLD PATENT DOCUMENTATION*, 5(2): 60-7, 1983.
2. BORKO, H. & BERNIER, C.L. Indexing concepts and methods. New York, Academic Press, 1978.
3. FLORIO, J.P. COLLEAGUE: an evolving medical information network. *INTERNATIONAL JOURNAL OF MICROGRAPHICS AND VIDEO TECHNOLOGY*, 4 (1):13-6, 1985.
4. HARMAN, J. Reuters: a survey of end user searching. *ASLIB PROCEEDINGS*; 38 (1): 35-42, 1986.
5. HUTCHINS, W.J. Language of indexing and classification. England, Peter Peregrinus, 1975.
6. KRUGER, K.J. MARC tags and retrospective conversion: the editing process. *INFORMATION TECHNOLOGY AND LIBRARIES*, 4 (1): 53-57, 1985.
7. LANCASTER, F.W. Toward paperless information systems. New York, Academic Press, 1978.
8. LOVE, R.A. & GARSON, L.R. Precision in searching the full-text database - ACS journals on-line. In: *NATIONAL ON-LINE MEETING PROCEEDINGS*, New York, April 30-may 2. 1985.

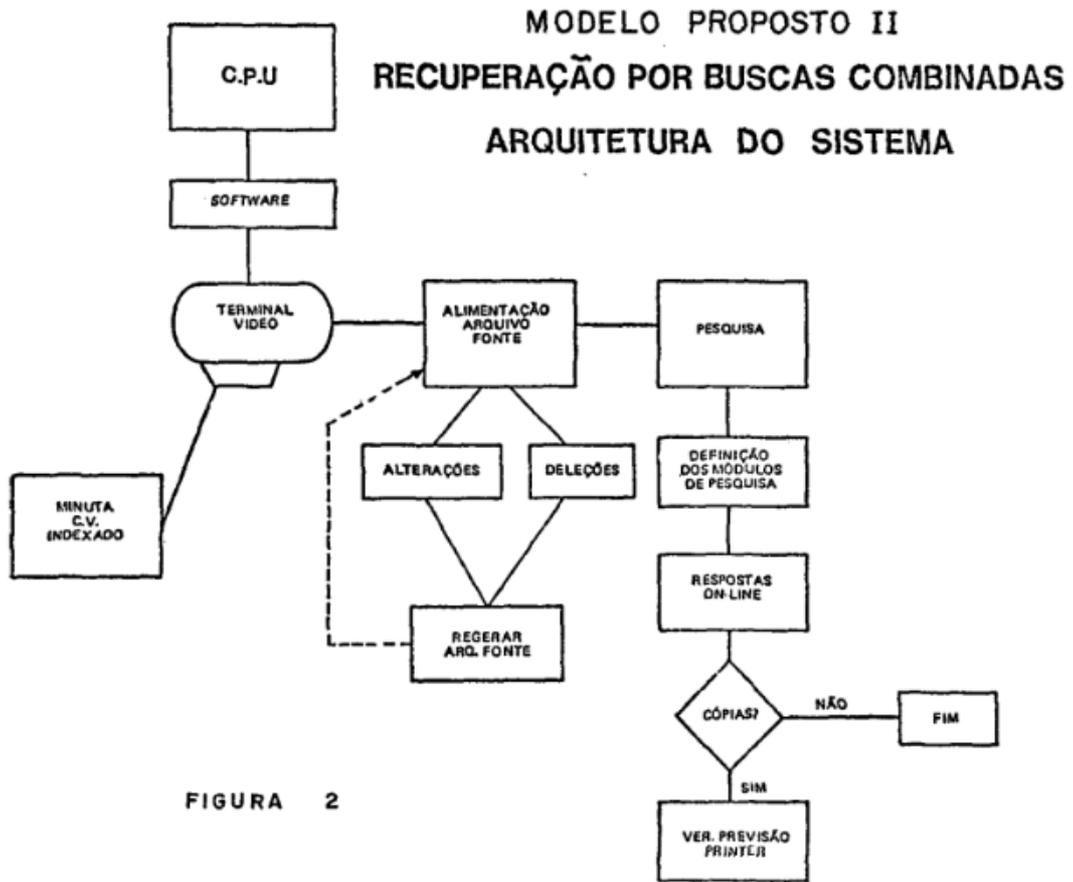


FIGURA 2