

Sistema de impressão para bases bibliográficas: características desejáveis

Fernando Moreira da Silva Martins Pereira

Jayme Leiro Vilan Filho

Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT

SCN Quadra 02 – Bloco K

70710 Brasília-DF

Resumo – A necessidade crescente de produtos impressos, diversificados e elaborados, determina o surgimento de sistemas de impressão de modo a atender a essas exigências. Esses sistemas devem ser fáceis de manusear e genéricos, permitindo, ao usuário, criar novos produtos de uma maneira natural. A confecção de produtos impressos, através de computadores, depende da estrutura de dados da base, do formato bibliográfico adotado e também dos recursos técnico-administrativos à disposição da entidade. O processo de execução de um produto deve ser dividido em três etapas: seleção, ordenação e impressão final, sendo que as características que cada uma deve apresentar depende do grau de elaboração desejado nos produtos impressos.

Palavras-chave – Bases de Dados Bibliográficos; Saídas; Referência Bibliográfica; Índice; Bibliografia.

1 Introdução

A obtenção de produtos impressos foi, no decorrer das décadas de 50 a 70, a grande motivação para a criação de bases bibliográficas e o desenvolvimento de sistemas de impressão. Com a evolução das técnicas de recuperação e comunicação de dados, a obtenção de produtos impressos passou a dividir essa importância com a busca retrospectiva e outras formas de disseminação de dados em linha.

Entretanto a impressão ainda é, e continuará por muito tempo, sendo um fator importante da exploração de bases de dados. No decorrer de todos esses anos surgiram inúmeros sistemas de impressão que, de um modo ou de outro, atenderam as necessidades de cada base.

A maioria desses sistemas tem como características básicas a inflexibilidade em relação aos formatos, a difícil manipulação pelo usuário e um pe-

queno número de variações frente a crescente complexidade dos produtos impressos. Já existem, principalmente no exterior, sistemas bastante eficazes e específicos para a elaboração de produtos impressos a partir de bases bibliográficas. No entanto, a crescente complexidade dos formatos de impressão exigem cada vez maior eficiência desses sistemas, tornando obrigatório o desenvolvimento de um número cada vez maior de opções.

De um modo geral, o artigo procura mostrar quais os componentes de um sistema de impressão para bases bibliográficas estabelecendo suas características desejáveis.

2 Aspectos Relevantes

Em primeiro lugar é importante lembrar que os produtos são fortemente dependentes da estrutura de dados da base. Agrupar ou separar dados em campos não é uma questão subjetiva. Essa decisão determina as possibilidades de impressão, tornando o sistema mais simples ou mais complexo de acordo com a estrutura adotada e, por vezes, inviabilizando a geração dos produtos desejados. A complexidade do sistema é proporcional à complexidade dos produtos, ou seja, quanto maior for o número de detalhes dos produtos, mais complexo é o sistema. A estrutura dos dados e o formato do conteúdo devem ser pensados não só para atenderem as necessidades de produtos hoje existentes, mas devem ser voltados para as necessidades que deverão surgir a médio e longo prazo. Isto porque, com o passar do tempo, o usuário tende a exigir mais alternativas de impressão, não se satisfazendo mais apenas com aquelas existentes. Nesse momento, deve ser levado em consideração se o sistema deve ser genérico ou específico. Um sistema específico apresenta resultados rápidos, porém temporários. Com o aumento das necessidades do usuário, esse sistema torna-se fator limitante das atividades desenvolvidas, pois não permite a geração de produtos diferentes daqueles previstos inicialmente. Já os sistemas genéricos requerem um pouco mais de tempo para serem confeccionados, mas tem um maior tempo de vida, permitindo ao usuário aumentar a complexidade de impressão de maneira natural.

Outra questão a ser levantada antes da decisão de desenvolver um sistema de impressão é a da administração da base de dados e dos produtos, o que demanda recursos humanos qualificados. É preciso pensar em três pontos: na periodicidade de geração dos produtos, nas correções necessárias para emití-los e na definição dos formatos de impressão. A periodicidade é importante para satisfação do público da base de dados. Bases cujas informações tornam-se rapidamente desatualizadas ou perdem o seu maior valor, como uma base de eventos (congressos, seminários, etc), devem gerar produtos com a maior frequência possível. Além disso a periodicidade exige

uma previsão de entrada de novos dados, o que requer alocação de pessoal. As correções de última hora são necessárias, pois no momento da impressão sempre surgem dados incorretos e algumas normalizações devem ser executadas. Outro ponto importante, é a necessidade de pessoal qualificado para definir os formatos de saída, seguindo normas e escolhendo a maneira mais correta de apresentá-los.

Os recursos, tanto humanos quanto computacionais, determinam as possibilidades de um sistema de impressão. Para evitar frustrações, o sistema deve se adequar às possibilidades organizacionais, sem no entanto, significar a adoção de soluções simplórias que geram insatisfações rapidamente.

Após a exposição dos passos a serem observados no início de um sistema de impressão, passaremos as características desejáveis.

3 Características Desejáveis

Um sistema de impressão deve ser de fácil utilização, para permitir que o próprio usuário final possa definir os seus produtos. Essa facilidade deve estar associada a flexibilidade de criar e alterar todas as características de impressão evitando ao máximo, ou até mesmo eliminando, as características pré-estabelecidas e imutáveis.

Esse sistema deve permitir a geração de relatórios dependentes entre si, no caso de uma bibliografia (corpo e índices associados), ou não, como um conjunto de relatórios operacionais gerados a partir de um mesmo grupo de registros a serem impressos.

É interessante que o usuário possa realizar testes de impressão no próprio terminal de vídeo, evitando desperdício de tempo e papel. Caso não exista uma impressora à disposição do usuário, permanentemente, o fator tempo torna-se mais crítico, pois o usuário deverá esperar pela liberação da impressora para poder iniciar seus testes.

Antes de continuarmos a descrever as características desejáveis, cabe definir que a emissão é composta de três etapas: seleção, ordenação e impressão. A seleção é a etapa de escolha dos registros (referências) que deverão ser impressas. A ordenação é a etapa na qual os registros selecionados são classificados segundo chaves de ordenação. Na etapa de impressão são importantes as características de apresentação do relatório, como cabeçalho, paginação, número de colunas, etc. Cada etapa deve ser definida em separado permitindo aproveitamento posterior em outras emissões.

Definida a composição de um relatório, é importante que o sistema permita ao usuário solicitar a reexecução apenas da etapa que por erro de definição gerou um produto não desejado. Essa possibilidade de reexecução em separado de cada uma das três etapas, evita a necessidade do reprocessamento das etapas que estão corretas. Seja, por exemplo, um relatório em que o número de colunas de texto desejado é dois e que por algum engano na definição ficou somente com uma coluna. Ao verificar esse erro, bastaria ao usuário alterar essa característica e solicitar a reexecução apenas da etapa de impressão. Nesse exemplo não seriam reprocessadas as etapas de seleção e ordenação que já estavam corretas, propiciando uma economia de tempo na geração final do produto. Outro fator que permite a redução de tempo é o reinício. Sem ele, a cada interrupção do computador, é necessário refazer toda a geração do produto. Essas características são muito importantes se considerarmos que um produto pode levar horas em processamento.

É importante que o sistema permita a reimpressão de uma faixa de páginas para repor as que apresentarem algum tipo de falha como a fita de impressão soltar ou desalinhamento do mecanismo da impressora.

Existem duas opções quanto a geração de produtos a nível de projeto computacional. O produto pode ser impresso diretamente, ou gerado em disco magnético e posteriormente impresso. Na primeira opção os programas de computador serão mais complexos e usuário só terá aquela cópia do produto. Caso seja necessário tirar uma nova cópia é preciso reprocessar toda a impressão. No caso da geração em disco, o produto pode ser reimpresso rapidamente, pois a geração de produto já se encontra pronta bastando apenas reimprimir. Além disso, a emissão fica salva como foi impressa não importando que novos dados tenham entrado na base. Por fim, permite que a impressão possa ser transferida de um equipamento para outro (microcomputador para minicomputador e vice-versa) ou para fica magnética. A escolha de um dos dois tipos depende do recurso computacional disponível, como por exemplo, área em disco magnético para permitir o armazenamento da emissão.

Estabelecidas as características gerais que um sistema de impressão deve ter, passaremos a detalhar as características desejáveis em cada uma das três etapas que compõem um relatório: seleção, ordenação e impressão.

3.1 Seleção dos Registros

O sistema deve permitir duas formas de seleção: através de um recuperador de informações ou por faixa de registros. No recuperador os registros são selecionados através de buscas apoiadas em listas invertidas, o que tornam a busca mais rápida, ou diretamente nos registros (busca textual). Na seleção por faixa de registros, o usuário indica que registros participam da

impressão e dentro dessa faixa pode realizar uma seleção segundo um grupo de caracteres ou faixa de valores.

Outra característica na seleção é imprimir mais de um relatório em conjunto, usando a mesma seleção de registros, mesmo que os relatórios sejam independentes na ordenação e impressão.

3.2 Ordenação dos Registros Seleccionados

Na ordenação são indicadas chaves onde cada uma pode ser formada por um ou mais campos de dados. Consideramos que o número máximo de chaves por relatório está entre 5 e 7. A composição de cada chave deve possibilitar a concatenação de campos para gerar um valor de chave, bem como a extração de valores de chave de campos distintos, sem concatenação, para compor uma mesma chave de ordenação. Como exemplo do primeiro tipo poderia ser obtido valor de chave a partir da concatenação do campo de título e subtítulo. No segundo tipo poderiam ser extraídos valores de chave a partir de palavras do conteúdo de campos como assunto e termos livres sem concatenação.

Cada chave de ordenação deve permitir diferentes formas de extração dos seus valores: por palavra, onde cada vocábulo extraído de um conteúdo de campo gera um valor de chave; por conteúdo inteiro, onde cada ocorrência de campo gera um valor de chave; por delimitador, onde cada conjunto de caracteres que estiver entre um ou dois delimitadores gera um valor de chave; por rotina específica, onde o analista de sistemas gera uma rotina que realize a extração desejada pelo usuário. Essa última forma deve ser encarada como uma abertura deixada para que o usuário possa suprir uma necessidade particular de impressão.

Além disso, o usuário deve poder solicitar que sejam ignorados caracteres do conteúdo de campo antes da sua ordenação, assim como, indicar que as chaves sejam colocadas em caracteres maiúsculos (normalizadas) para evitar a desordenação proveniente do preenchimento do campo com letras maiúsculas e minúsculas. Também deve possibilitar ao usuário estabelecer quais conjuntos de caracteres não devem ser considerados como chave usando para isso uma tabela de termos proibidos. Se um registro não gerar valor de chave, deve ser possível a colocação do caracter branco, ao invés de ignorar esse registro na ordenação. Outra possibilidade é a de decodificar o valor de chave extraído usando uma tabela. Dessa maneira o valor extraído é transformado no valor correspondente na tabela, para então ser ordenado.

Devem existir três formas de ordenação: ascendente, onde os valores de chave são ordenados do menor valor para o maior; descendente, onde os

valores de chave são ordenados do maior valor para o menor; hierárquico, onde a ordenação dos valores seguem uma ordem predefinida em uma tabela hierárquica, como a tabela de Áreas do Conhecimento do CNPq.

3.3 Impressão do Produto

Nessa etapa deve ser permitido colocar cabeçalho e rodapé em cada página dos relatórios, possibilitando a colocação de data, numeração de página alternada na esquerda e direita se for par ou ímpar, respectivamente. Por vezes, os produtos são continuação de produtos que já foram publicados ou é necessária a colocação de páginas de apresentação. Dessa maneira, a página inicial de um produto deve ser parâmetro de impressão, para permitir essas alternativas.

Outros parâmetros importantes são o número de colunas de texto, número de caracteres por coluna, espaço entre as colunas, linhas em branco entre cada linha impressa, linhas em branco antes de imprimir cada referência e cada chave de ordenação, margem esquerda e o número de linhas da mancha, que cada relatório deve ter. Com esses parâmetros a impressão pode aproveitar melhor a página e a disposição das informações pode ficar mais atraente.

Para evitar a quebra indevida de referências, ou separação da chave de ordenação de sua referência correspondente ao final de coluna ou página, deve haver um parâmetro que denominamos tolerância. A tolerância é o número mínimo de linhas disponíveis antes do fim da coluna (ou página), para que o sistema inicie a impressão de um conjunto de chaves e a referência correspondente.

No caso de relatórios que se relacionam, como por exemplo, uma bibliografia na qual o corpo está relacionado com seus índices através da numeração das referências, é necessário que exista um parâmetro que indique qual relatório será o principal e quais serão os relatórios associados a ele, e que essa numeração seja impressa de modo a ligar cada termo do índice a um número de referência no corpo.

Quanto à impressão das chaves de ordenação, devem existir algumas opções básicas: não imprimir a chave, quando se deseja apenas classificar as referências sem imprimir a chave de ordenação; decodificar um valor de chave segundo uma tabela antes da impressão; imprimir o valor de chave como foi ordenado; imprimir o valor de chave como se encontra armazenado na base; colocação de alguns caracteres do valor de chave em caracteres maiúsculos e outros minúsculos, permitindo uma melhor qualidade na apresentação do relatório; determinar o posicionamento da chave em qualquer ponto da linha.

A impressão das referências deve permitir que o usuário defina: quais campos de dados devem ser impressos; quais literais devem separá-los; se estes literais são condicionados ou não à existência de conteúdo de campo a cada referência; posição inicial de impressão de um campo diferenciada para a primeira linha e as demais (identação); imprimir parte de um conteúdo de campo; espaçamento vertical (salto de linha) e horizontal (tabulação e espaço em branco); impressão de parte ou todo o conteúdo de campo em letras maiúsculas; imprimir o número da referência na base de dados; imprimir identificador de campo; permitir a execução de um conjunto de ações condicionadas a um conjunto de fatores; permitir a impressão de dados de uma ou mais bases em uma mesma referência; e rotina específica para permitir um tratamento particular.

Por fim, o usuário deve poder escolher a forma de separação de palavras, a colocação de chaves de ordenação e parte das referências em destaque. Na separação de palavras, existem três formas de tratamento possíveis: não separar palavras mantendo alinhamento à direita e à esquerda, alinhar somente à esquerda ou realizar separação silábica. Deve ser evitada a colocação de caracteres brancos para alinhar à direita em literais colocados entre conteúdos de campos, pois são normalmente parte de normas rígidas. O destaque pode ser dado colocando as informações em negrito, sublinhado, itálico e outros.

4 Conclusão

A geração de produtos impressos continua sendo a principal forma de disseminação de informações bibliográficas sendo portanto a maior motivação para a criação de sistemas de impressão.

Como em geral tanto as bases de dados bibliográficos como seus produtos envolvem uma complexidade crescente, os sistemas capazes de suportá-los são cada vez mais complexos. Essa complexidade não se dá apenas ao nível do sistema de impressão, mas também, a nível do sistema gerenciador da base, pois o produto impresso depende muito da estrutura dos dados que o compõem.

Em função disso, entidades que desejem desenvolver esses sistemas, devem primeiramente procurar idéias nos já existentes e auxílio técnico em organizações da área de informação. Além disso, é fundamental que a entidade saiba claramente que produtos deseja e disponha de toda a infra-estrutura técnico-administrativa para manipular bases de dados bibliográficos e seus produtos impressos.

Os sistemas de impressão devem ser genéricos e de fácil manuseio. Essa generalidade é um fator importante se levarmos em conta que um produto evolui a cada emissão e que, além disso, outros são criados no decorrer do tempo.

Como não existem regras rígidas para a determinação de um bom sistema de impressão, talvez por sua grande variedade de características, acreditamos que cada entidade deva procurar no mercado, ou desenvolver, seus sistemas de impressão de acordo com suas necessidades.