

BASES DE DADOS: METODOLOGIA PARA SELEÇÃO E COLETA DE DOCUMENTOS

Rogéria Fernandes Albrecht
Maria Lourdes Blatt Ohira

Resumo

Apresenta critérios de seleção e de coleta de documentos para ingresso em bases de dados, com o objetivo de oferecer uma metodologia a ser utilizada no planejamento de bases de dados.

Palavras-chave: Bases de Dados - Metodologia; Política de Seleção de Documentos; Política de Coleta de Documentos.

I INTRODUÇÃO

O conceito de bases de dados passou a ser identificado na literatura, aproximadamente nos anos 60, com o rótulo de Banco de Dados. Cianconi (1987) aponta grande confusão de conceitos e de terminologia na literatura especializada, que trata banco de dados e bases de dados como sinônimos e sugere conceituar **bases de dados** como “*um conjunto de dados interrelacionados, organizados de forma a permitir recuperação de informações*”, e **banco de dados** como “*um conjunto de bases de dados*”. Para Heemann (1997, p. 2) a literatura em Ciência da Informação, de maneira geral, trata as bases de dados atuais “*como um arquivo ou um conjunto de arquivos computacionais no qual são armazenados dados, permitindo a recuperação e atualização de informações*”.

Guinchat & Menou (1994, p. 295), consideram bases de dados diferentes de bancos de dados, assim conceituando-os: “*uma*

base de dados é um conjunto organizado de referências bibliográficas de documentos que se encontram armazenados fisicamente em vários locais (...)". Enquanto que, "os bancos de dados tratam das informações factuais, numéricas ou textuais diretamente utilizáveis", isto é, proporcionam acesso direto à informação primária. As bases de dados são armazenados por meios magnéticos e acessados local ou remotamente.

Ferreira (2000), mostra a evolução das bases de dados por ordem cronológica, de acordo com a data de surgimento das mesmas, a saber:

- 1951 - Bases de Dados Numéricos;
- 1960 - Bases de Dados Bibliográficos;
- 1970 - 10 Bases disponíveis em 2 Bancos de dados (acesso on line);
- 1986 - Bases em suporte de CD-ROM;
- 1990 - 3.200 Bases hospedadas em mais de 40 Bancos de Dados (acesso em rede: Dialog, Orbit , Questel e outras) e;
- 1997 - Milhões de computadores ligados em rede. - Globalização da informação.

De acordo com Cunha (1994) e Rowley (1994), as bases de dados podem ser classificadas em Bases de Dados Referenciais e Base de Dados de Fontes.

Bases de dados referenciais – "encaminham ou orientam o usuário para outra fonte, que pode ser um documento ou uma instituição ou um indivíduo, a fim de obter informações adicionais ou conseguir o texto integral de um documento" (Rowley, 1994, p. 68). Abrangem, tipos diversificados de bases de dados como:

Bases de Dados Bibliográficos - Incluem citações bibliográficas acompanhadas ou não dos resumos dos trabalhos;

Bases de Dados Catalográficos - Representam o acervo de uma biblioteca ou de uma rede de bibliotecas, sem indicação do conteúdo dos documentos.

Bases de Dados Referenciais – que referenciam informações ou dados, como nomes e endereços de instituições, dados característicos de guias, cadastros, diretórios etc

Bases de Dados de Fontes - contém os “*dados originais e constituem um tipo de documento eletrônico*”. Podem ser agrupados segundo seu conteúdo em:

Bases de Dados Numéricos - Incluem dados numéricos e estatísticos.

Bases de Dados de Texto Completo - contém notícias de jornal, especificações técnicas, artigos de periódicos, dicionários.

Bases de Dados Textuais e Numéricos - mistura de dados textuais e numéricos.

Bases de Dados Gráficos - apresentam fórmulas químicas, imagens, logotipos.

Em relação ao **conteúdo** das bases de dados, as mesmas podem ser orientadas por disciplina, por missão, por problemas, ou ainda, podem ser de caráter multidisciplinar ou interdisciplinar (Martyn, 1985 apud Sayão, 199-):

Por disciplina – Bases e dados que refletem o corpo de conhecimento de uma determinada área.

Por Missão – São bases de dados que têm como objetivo apoiar missões bem determinadas.

Por Problema – Bases de dados orientadas para a solução de problemas específicos

Interdisciplinar – Bases de dados que cobrem mais de uma área de conhecimento.

Multidisciplinar – Bases de dados que cobrem uma grande diversidade de áreas do conhecimento.

O **objetivo** de uma base de dados é fornecer informação atualizada (recursos estruturais), precisa e confiável (não dar a informação pela metade) e de acordo com a demanda (oferecer o que o usuário necessita). Um requisito essencial para uma base de dados

é não somente a armazenagem eficiente de dados, mas também o fornecimento de mecanismos eficientes de recuperação.

A seleção e coleta adequada de documentos para ingresso em bases de dados, deve seguir uma política definida, visando a obtenção de um maior grau de consistência e eficiência em redes e sistemas de informação. Através de um conjunto de diretrizes e normas, objetiva estabelecer adoções, delinear estratégias gerais, determinar instrumentos e delimitar critérios que facilitem a tomada de decisão na composição e desenvolvimento de coleções, de acordo com os objetivos e os usuários de um sistema de informação. Esta política deve ser flexível e atualizada, e expressa de forma a facilitar as decisões e a justificar a utilização ou não, de determinados itens (Chagas, 1988).

2 METODOLOGIA

Com o objetivo de fornecer uma metodologia que poderá ser utilizada no planejamento e definição de bases de dados, no que tange as atividades de seleção e coleta de documentos para inclusão nas mesmas, serão analisados os critérios e procedimentos adotados pelos seguintes autores: Dodebei (1986) ; Chagas (1988) ; Ohira (1994) ; Lambert (1996) ; Lambert & Barreiro (1996) e (Sayão, 199-) que compreende os seguintes passos:

2.1 Tipo de Literatura

O primeiro aspecto a ser analisado na coleta de documentos, refere-se a seleção da literatura, obedecendo alguns critérios básicos. Dentre as inúmeras classificações adotadas para os diversos tipos de documentos, optamos por Población et al. (1996) que agrupa em literatura cinzenta ou não convencional e literatura branca ou convencional. Para os autores se por um lado, a **literatura cinzenta** caracteriza-se pela sua tiragem reduzida, de divulgação restrita e conseqüentemente pela dificuldade de acesso, por outro lado, ela flui com rapidez entre os pares, pelo fato de ser **não convencional**, isto é,

independe da formalização exigida pela apresentação dos **documentos convencionais** – livros e publicações periódicas – que integram o grupo da chamada **literatura branca**.

De acordo com depoimentos obtidos por Foskett & Hill apud Población et al. (1996), observa-se que, *90 % das informações de que os pesquisadores necessitam são provenientes da literatura cinzenta, que não possui um controle bibliográfico, como ocorre com a literatura branca, uma vez que os artigos de periódicos são indexados em bases de dados nacionais e internacionais das diversas áreas do conhecimento.*

2.1.1 Documentos Convencionais

Os diversos tipos de documentos convencionais estão assim divididos:

Livro: obra literária, científica ou artística que compõe, em regra, um volume.

Capítulo de Livro: parte de um livro, podendo ser escrito pelo próprio autor do livro e ou com a colaboração de outros autores.

Artigo de Periódico: trabalho técnico científico ou cultural que visa principalmente maior agilidade na divulgação do assunto tratado, sendo inseridos em publicações editadas em fascículos, números ou partes, com intervalos regulares, ou não, com título comum, plano definido, assuntos e autores diversos. Nos artigos de periódicos, o assunto deve ser original, e redigido de acordo com os padrões científicos exigidos pela ciência, permitindo que os pares da comunidade possam avaliar as observações, repetir os experimentos e julgar os processos intelectuais utilizados.

Artigo de Jornal: publicação com o objetivo de transmitir informações gerais sobre determinado assunto, com caráter mais informativo do que científico, veiculado em periódicos diários, semanais e mesmo de periodicidade mensal.

2.1.2 Documentos não-convencionais

A produção de literatura não convencional é bastante grande, principalmente nos países do Terceiro Mundo, necessitando assim de ações diferenciadas no processo de coleta. Os documentos não-convencionais agrupam-se da seguinte forma:

Comunicação em Evento: trabalhos apresentados em eventos (congressos, seminários e similares), sob a forma de comunicações em painel ou poster, sendo o texto, na íntegra e/ou os resumos publicados nos respectivos anais.

Tese: é uma proposição, baseada numa hipótese, que se caracteriza por uma nova perspectiva de um tema já apresentado ou apresentação de um tema original.

Dissertação: é o resultado de um trabalho de pesquisa, consiste no estudo, o mais completo possível, do tema escolhido, expressando conhecimentos do autor a respeito do assunto e sua capacidade de sistematização.

Relatório: os relatórios devem expor de forma sistemática o assunto de que tratam. De acordo com o objetivo, são classificados como: relatórios técnico - científicos - documento original pelo qual se faz a difusão da informação corrente, sendo ainda o registro permanente das informações obtidas. Sua elaboração serve, principalmente, para descrever experiências, investigações, processos, métodos e análises; relatórios de viagem - é o documento por meio do qual são fornecidas informações sobre viagem realizada, indicando data, destino, duração, participantes, objetivos e atividades desenvolvidas; relatórios de estágios e visitas - visa a descrição do local onde foi realizado o estágio ou a visita, o período de duração, as atividades desenvolvidas pelo estagiário, ou as observações feitas pelo visitante; relatório administrativo - é a comunicação escrita submetida à apreciação de uma autoridade superior, geralmente ao término de um exercício, relatando a atuação administrativa; relatórios para fins especiais - documento organizado de forma particular, que especifica instruções para otimizar o uso de materiais, máquinas, dispositivos e equipamentos.

Outros exemplos de relatórios para fins especiais são: levantamento de produção, orçamento de pesquisas, registro de patentes, manuais de software.

Monografia ou Trabalho Acadêmico: monografia é uma exposição exaustiva de um problema ou assunto específico, investigado cientificamente. Em algumas instituições de ensino, monografia é o trabalho de pesquisa apresentado como requisito para conclusão de curso. Também denominada de trabalho acadêmico, consiste em exposições por escrito sobre temas atribuídos em disciplinas de cursos de graduação ou pós-graduação nos diversos níveis.

2.2 Abrangência do Assunto e/ou Área Temática

Uma das soluções para o problema de definição do assunto, é iniciar o trabalho com um escopo mais reduzido e durante sua execução, com a experiência de lidar com o tema, acrescentar, gradativamente, novos assuntos (Lambert & Barreiro, 1996).

Alguns assuntos são muito abrangentes, necessitando de conceitos que pertencem a diversas áreas de conhecimento, o que pode inicialmente, gerar dificuldades. Desta forma, um número limitado de assuntos, no começo, e o acréscimo de novos, durante o trabalho, facilitarão a inclusão de novos temas. Devido a evolução do conhecimento e ao dinamismo da base de dados, deve-se estar atento às alterações necessárias com relação a inclusão e exclusão de determinados assuntos.

2.3 Limite Geográfico

Outro aspecto a considerar refere-se a delimitação do limite geográfico, no sentido de coletar somente os documentos cuja abrangência tenha sido definida anteriormente: local, regional, estadual, nacional e/ou internacional. Para Lambert & Barreiro (1996, p. 6), *o limite geográfico vai determinar as fronteiras do trabalho de coleta, vai definir as fontes de informação que serão*

utilizadas, se vão ser limitadas às fontes nacionais ou se serão direcionadas também às fontes internacionais.

2.4 Limite de Tempo

Pelo fato de que, o volume da literatura publicada na área técnico-científica, ser muito grande, na maioria das vezes, a delimitação do tempo é fundamental para o interesse ou não da inclusão dos documentos na base de dados. No caso de estabelecimento de períodos de tempo que retroagem muitos anos, poderá haver complicações na fase de coleta dos documentos, causando duplo trabalho, com a coleta dos documentos antigos, acumulados por muitos anos, e os que vem sendo publicados (Lambert & Barreiro, 1996).

2.5 Identificação das Fontes de Informações

Tendo em vista conhecer o que é produzido em termos de literatura técnico-científica sobre determinado assunto, dentro do período de abrangência e limites geográficos estabelecidos, deve-se ter, inicialmente, o conhecimento das fontes produtoras. Configurada esta necessidade, sabe-se que ela se constitui no primeiro passo das ações para a formação da base de dados, já que esta é a forma mais eficaz de se obter a literatura técnico-científica por elas gerada. As ações que compreendem a coleta de documentos acontecem em dois estágios: localizar as fontes produtoras e analisar os documentos coletados, objetivando assim, a localização de outras fontes - documentos, pessoas e instituições (Dodebei, 1986, p. 317).

A identificação das **pessoas** como fontes de informações, leva de uma maneira direta ou indireta aos indivíduos que podem contribuir para a alimentação da base de dados e são representadas pelos autores, editores e orientadores de dissertações e teses. A identificação de pessoas pode levar a identificação de instituições, de publicações e de outras pessoas.

Fontes de informações definidas como **instituições**, são aquelas que editam ou são depositárias de documentos de interesse

da base de dados. Estão incluídas as bibliotecas, as instituições de pesquisa, as instituições públicas e privadas, entre outras a serem identificadas. A identificação de instituições pode levar a identificação de pessoas, de publicações e de outras instituições.

As **publicações** enquanto fontes de informações, estão representadas pela literatura técnico-científica impressa em papel, microfilmada, armazenada em suporte magnético ou em CD-ROM se constituindo em: periódicos, relatórios, bibliografias, entre outros (Lambert, 1996).

No trabalho de levantamento das fontes de informação ocorrem dois fatos importantes: a possibilidade de se conseguir a partir de um documento, indicações de novas fontes de informações – no caso de bibliografias, na lista de referências bibliográficas de um documento, teremos, uma série de novas referências que nos levarão a outras referências, formando assim, a cadeia de fontes de informações. Este desdobramento é positivo, uma vez que fornece novo e farto material de trabalho, sendo o exame das listas de referências bibliográficas, o ponto fundamental da metodologia de coleta de documentos. O segundo ponto importante neste contexto é o de surgirem em algum ponto desta cadeia, fontes de informação já identificadas e coletadas, ou seja, repetidas (Lambert & Barreiro, 1996).

2.6 Localização das Fontes de Informação

A etapa seguinte à identificação das fontes de informações, é a localização destas fontes. Segundo sua origem as fontes podem ser primárias, secundárias e terciárias, definidas por Lambert (1996), como:

Fontes primárias – são aquelas que contém ou divulgam informações originais ou que apresentam, sob forma original, informações já conhecidas. As fontes primárias são as mais importantes, por representarem a grande produção técnica e científica da área. Nelas incluem-se: os livros, os periódicos e publicações

seriadas, os preprints e anais de eventos, os relatórios técnicos, as normas técnicas, as teses e dissertações e as patentes.

Fontes secundárias – são as que organizam, sob forma de índices e resumos, as informações de fontes primárias, facilitando assim o conhecimento e o acesso às mesmas. As publicações englobadas nesta categoria, normalmente são designadas como “obras de referência”. São elas: as bibliografias, os handbooks ou manuais técnicos, os boletins de bibliotecas, os informativos de associações e sociedades de classe, os catálogos de teses e de produção científica.

Fontes terciárias – são as que orientam o usuário para a utilização das fontes secundárias e primárias, facilitando a localização e o acesso às informações. Elas representam o ponto de partida para as ações da coleta. Constituem-se dos seguintes documentos: banco de bibliografias ou bibliografias de bibliografias, os catálogos de editores, os catálogos coletivos de eventos, de normas técnicas, de periódicos, os CD-Rom que informam sobre acervos de bibliotecas, os calendários de eventos e outros.

2.7 Organização e definição dos contatos

A organização dos contatos é uma fase muito importante para se proceder a coleta de documentos. Cada pessoa e instituição pode ser abordada de diferentes formas. Estes contatos podem ser feitos por: correspondências formais - cartas-padrão ou cartas-circulares (com redação clara e objetiva, deve conter informações sobre o solicitante, a finalidade do pedido, bem como informar sobre o funcionamento do sistema, as características da base de dados e a abrangência do assunto); formulários - padrão de pedidos entre bibliotecas (devem constar informações sobre o título do trabalho solicitado, autor, ano, volume, número, paginação, total de páginas do artigo; formulários do COMUT – existente no Brasil através do sistema de comutação bibliográfica, é um programa que tem uma rede de bibliotecas, espalhadas em vários locais do país, também poderá ser utilizado um formulário padrão, facilitando o contato e a

obtenção de cópias. Poderão ser feitos também através de contatos informais – telefone, fax, e-mail, etc.

Muito importante são os contatos pessoais, uma vez que, permitem fazer um trabalho de conscientização para pesquisadores e/ou instituições sobre as necessidades de registro do trabalho técnico-científico e da cooperação, bem como, através do contato pessoal, poderá haver a possibilidade da geração de outras fontes de referência (Lambert, 1996) ; (Lambert & Barreiro, 1996) e (Dodebei, 1986).

2.8 Feedback aos Autores e Instituições-Fonte

O feedback ou retorno, constitui da última etapa do trabalho no sistema de coleta. Após a inclusão do autor na base de dados, este deverá ser referenciado na bibliografia impressa e comunicado através de correspondência da inclusão de seu trabalho na base de dados, como forma de agradecimento pela colaboração e como incentivo ao envio de futuros trabalhos de sua autoria ou de seus pares (Lambert, 1996).

Segundo Lambert & Barreiro (1996), o feedback:

“É uma etapa fundamental para se estabelecer um fluxo constante de documentos. Muitas vezes, a falta de vontade em remeter ou emprestar um documento, se deve ao fato de que as pessoas não sabem o que vai ser feito daquele documento. Para evitar que isto aconteça basta informar ao autor do trabalho, ou à pessoa que enviou o exemplar do documento, quando ele foi incorporado à base de dados. Ainda melhor é enviar uma cópia da página da bibliografia (caso haja) que contenha a referência bibliográfica em questão.”

3 CONCLUSÃO

A importância da adoção de uma política de seleção para ingresso de documentos em bases de dados, independente da área de

conhecimento a que pertença, é fundamental para se obter a qualidade e eficácia do produto que será disponibilizado ao público.

O processo de seleção e coleta de documentos deve ser constantemente realimentado, num ciclo interligado de atividades sistematizadas. Os critérios e procedimentos enumerados neste artigo, deverão ser analisados e usados pelos profissionais envolvidos com o planejamento e definição de bases de dados, com flexibilidade, afim de se adequarem às necessidades, decisões e justificarem a inclusão ou não de determinados itens, levando-se sem consideração os diversos tipos de documentos.

4 REFERÊNCIAS BIBLIOGRÁFICAS

CHAGAS, Maria Helena Simões de Carvalho. *Política de seleção de publicações para Bases de Dados*. Brasília : MINTER, 1988. (Trabalho apresentado na 2ª Reunião de Trabalho da Rede Minter de Informação Documentária – REMIDO, Brasília, 1988.) (mimeografado).

CIANCONI, Regina. Banco de Dados de acesso público. *Ciência da Informação*, Brasília, v. 16, n. 1, p. 53-59, jan./jun. 1987.

CUNHA, Murilo Bastos da. As tecnologias de informação e a integração das bibliotecas brasileiras. *Ciência da Informação*, Brasília, v. 23, n. 2, p. 182-189, maio/ago. 1994.

DODEBEI, Vera Lúcia Doyle Louzada de Mattos. Metodologia de coleta de documentos para bases de dados bibliográficos. *Revista de Biblioteconomia de Brasília*, Brasília, v. 14, n. 2, p. 312 - 327, jul./dez. 1986.

FERREIRA, Sueli Mara Soares Pinto. *Bases de dados*. Disponível em <<http://www.eca.usp.br/prof/sueli/cbd201/bases.html>> Acesso em: 5 set. 2000.

GUINCHAT, Claire , MENOU, Michel. *Introdução geral às ciências e técnicas da informação e documentação*. Trad. de Miriam Vieira da Cunha. Brasília : MCT/CNPq/IBICT, 1994, 540 p.

HEEMANN, Vivian. *Avaliação ergonômica de interfaces de bases de dados por meio de "checklist" especializado*. Florianópolis, 1997. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina.

LAMBERT, Maria Betânia Monte Alto. *Manual de coleta de Documentos para alimentação da base de dados do INIS*. Rio de Janeiro: Comissão Nacional de Energia Nuclear, 1996. 24 p.

LAMBERT, Maria Betânia Monte Alto. BARREIRO, Selma Chi. Metodologia de coleta de documentos para produção de bases de Dados bibliográficos. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 9, Curitiba; 1996. *Anais eletrônicos [disquetes]* Curitiba: UFPR, 1996.

OHIRA, Maria Lourdes Blatt. Biblioinfo : base de dados sobre automação em bibliotecas (Informática documentária) : 1986 -1994. *Ciência da Informação*, Brasília, v. 23, n. 3, p. 369-371, set./dez. 1994.

POBLACION, Dinah Aguiar, NORONHA, Daisy Pires, CURRÁS, Emilia. Literatura Cinzenta versus literatura branca: transição dos autores das comunicações dos Eventos para produtores de artigos. *Ciência da Informação*, Brasília, v. 25, n. 2, p. 228-242, maio/ago. 1996.

ROWLEY, Jennifer. *Informática para bibliotecas*. Trad. de Antônio Agenor Briquet de Lemos. Brasília : Briquet de lemos/Livros, 1994.

SAYÃO, Luis Fernando. *Produção e acesso à bases de dados bibliográficos*. (Curso Aplicação da tecnologia ao Desenvolvimento de Bibliotecas - Rede Bibliodata). Rio de Janeiro [199-]

DATA BASE: METHODOLOGY FOR DOCUMENTS SELECTION AND COLLECTION

Abstract

It presents selection criteria and document collection in order to be inserted in database so as to offer an approach which will be used in planning the latter.

Keywords: Data Base - Methodology; Document Selection Policy;
Document Collection Policy.

Rogéria Fernandes Albrecht - CRB 14/399

Especialista em Estratégias e Qualidade em Sistemas de Informação
Bibliotecária - Chefe: Biblioteca Pública Municipal Dom Daniel
Hostin, de Gaspar - Santa Catarina

Extraído da Monografia apresentada no Curso de Especialização em
Estratégias e Qualidade em Sistemas de Informação da Universidade
do Estado de Santa Catarina – UDESC

robiblio@zaz.com.br

Maria Lourdes Blatt Ohira - CRB 14/213

Bacharel em Biblioteconomia pela Universidade do Estado de Santa
Catarina – UDESC.

Mestre em Biblioteconomia pela Pontifícia Universidade Católica de
Campinas – PUCAMP

Professora do Curso de Biblioteconomia da Universidade do Estado
de Santa Catarina – UDESC

f2mlbh@pobox.udesc.br
