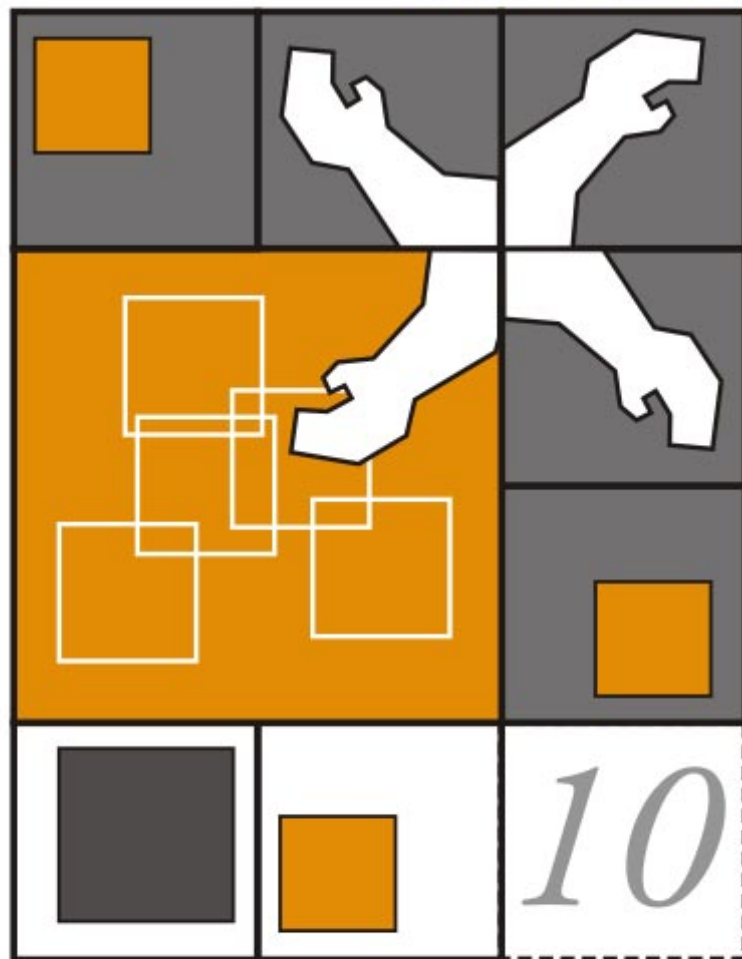


COMO FAZER

COMO ELABORAR  
VOCABULÁRIO CONTROLADO  
PARA APLICAÇÃO EM ARQUIVOS

Johanna Wilhelmina Smit  
Nair Yumiko Kobashi



**COMO ELABORAR  
VOCABULÁRIO CONTROLADO  
PARA APLICAÇÃO EM ARQUIVOS**  
COMO FAZER VOL. 10

Arquivo do Estado e Imprensa Oficial do Estado  
São Paulo  
2003

**COMO FAZER 10 - COMO ELABORAR VOCABULÁRIO CONTROLADO PARA APLICAÇÃO EM ARQUIVOS**

GOVERNO DO ESTADO DE SÃO PAULO  
Geraldo Alckmin  
GOVERNADOR

SECRETARIA DE ESTADO DA CULTURA  
Cláudia Costin  
SECRETÁRIA

DEPARTAMENTO DE MUSEUS E ARQUIVOS  
Marilda Suyama Tegg  
DIRETORA

DIVISÃO DE ARQUIVO DO ESTADO  
Fausto Couto Sobrinho  
DIRETOR

IMPrensa OFICIAL DO ESTADO  
Hubert Alquéres  
DIRETOR PRESIDENTE

Luiz Carlos Frigerio  
DIRETOR VICE-PRESIDENTE

Teiji Tomioka  
DIRETOR INDUSTRIAL

Richard Vainberg  
DIRETOR FINANCEIRO E ADMINISTRATIVO

CEETEPS  
CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA  
Prof. Marcos Antônio Monteiro  
DIRETOR-SUPERINTENDENTE

Prof. Remo Alberto Fevorini  
VICE-DIRETOR-SUPERINTENDENTE

Prof.<sup>a</sup> Laura Laganá  
CHEFE DE GABINETE

Arquivo do Estado de São Paulo  
R. Voluntários da Pátria, 596 – Santana  
CEP: 02010-000 – São Paulo – SP  
Fone/Fax: (11) 6221-4785  
www.arquivoestado.sp.gov.br  
arquivoestado@sp.gov.br

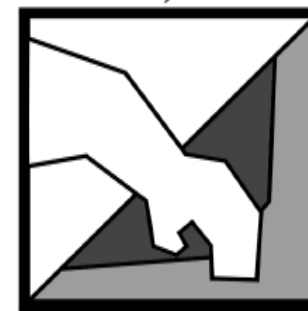
Imprensa Oficial do Estado  
Rua da Mooca, 1.921 – Mooca  
03103-902 – São Paulo – SP  
Tel.: (11) 6099-9800  
Fax: (11) 6692-9674  
www.imprensaoficial.com.br  
livros@imprensaoficial.com.br  
SAC 0800-123 401

Foi feito o depósito legal

JOHANNA WILHELMINA SMIT  
NAIR YUMIKO KOBASHI

**COMO ELABORAR  
VOCABULÁRIO CONTROLADO  
PARA APLICAÇÃO EM ARQUIVOS**  
COMO FAZER VOL. 10

Projeto



Como Fazer

ARQUIVO DO ESTADO/IMPrensa OFICIAL DO ESTADO  
SÃO PAULO  
2003

ARQUIVO DO ESTADO DE SÃO PAULO

Coordenação editorial  
LAURO ÁVILA PEREIRA  
MARIZA ROMERO

Editora responsável  
JULIANA PADUA MELO ALKMIN

Assistente editorial  
FERNANDO F. DE SOUSA LIMA

Revisão dos textos  
ANA PAULA DE FREITAS SANCHEZ  
LÍLIA RAMADAN VERÍSSIMO DE LIMA

Capa  
FERNANDO F. DE SOUSA LIMA

Diagramação  
PAULO F. M. DE BORGIA

Criação de arte  
TEREZA REGINA CORDIDO



ARQ/SP  
ASSOCIAÇÃO DE ARQUIVISTAS  
DE SÃO PAULO

Diretora  
HELOÍSA LIBERALLI BELLOTTO

Vice-Diretora  
ANA MARIA DE ALMEIDA CAMARGO

Secretária  
SILVANA GOULART FRANÇA GUIMARÃES

Tesoureira  
SILVIA COELHO HERNANDES

Coordenadora de cursos  
IEDA PIMENTA BERNARDES

Fotolito, impressão e acabamento  
IMPrensa OFICIAL DO ESTADO

Co-edição  
ARQUIVO DO ESTADO DE SÃO PAULO  
IMPrensa OFICIAL DO ESTADO

Apoio Técnico:  
CEETEPS - Centro Estadual de Educação Tecnológica Paula Souza  
Governo do Estado de São Paulo

S645C SMIT, JOHANNA WILHELMINA  
COMO ELABORAR VOCABULÁRIO CONTROLADO PARA APLICAÇÃO EM ARQUIVOS / JOHANNA WILHELMINA  
SMIT E NAIR YUMIKO KOBASHI. SÃO PAULO: ARQUIVO DO ESTADO, IMPrensa OFICIAL, 2003.  
56 P. (PROJETO COMO FAZER, 10).

BIBLIOGRAFIA  
ISBN: 85-86726-52-4 (ARQUIVO DO ESTADO)  
ISBN: 85-7060-216-2 (IMPrensa OFICIAL DO ESTADO DE SÃO PAULO)

1. ARQUIVOS – VOCABULÁRIO CONTROLADO 2. VOCABULÁRIO CONTROLADO 3. ARQUIVÍSTICA  
4. TESAUROS I. KOBASHI, NAIR YUMIKO. II. TÍTULO. III. SÉRIE.

CDD 025.017  
CDU 930.251

## SUMÁRIO

<b>APRESENTAÇÃO</b>	07
<b>SOBRE AS AUTORAS</b>	09
<b>O QUE É CONTROLE DE VOCABULÁRIO</b>	11
1.1. Controle de vocabulário e uso de vocabulário controlado	18
1.2. Variáveis intervenientes na construção de um vocabulário controlado para arquivos	20
<b>COMO FAZER O CONTROLE DE VOCABULÁRIO: PROCEDIMENTOS “MICRO”</b>	22
2.1. Normalização gramatical	22
2.2. Opções de grafia	23
2.3. Alterações nos nomes de pessoas e topônimos	24
2.4. Controle de sinonímia, homonímia e quase sinonímia	25
2.5. Adoção de termos compostos	27
2.6. Introdução de notas de escopo e notas de uso	30
2.7. Síntese dos procedimentos ‘micro’ de controle de vocabulário	33
<b>COMO FAZER O CONTROLE DE VOCABULÁRIO: PROCEDIMENTOS “MACRO” – ORDENAÇÃO DOS TERMOS</b>	34
3.1. Tipologia de vocabulários controlados	35
3.2. A construção de um plano de classificação	43
3.3. A construção de um tesauro funcional	44
3.4. A elaboração de listas alfabéticas complementares à ordenação dos termos	46
<b>CONCLUSÃO</b>	49
<b>O QUE HÁ PARA LER</b>	51
<b>ANEXO</b>	55

## APRESENTAÇÃO

O Arquivo do Estado de São Paulo tem sob a sua tutela um rico acervo e um grupo técnico capacitado para o desenvolvimento de suas atividades, o que lhe permite assumir o papel de relevo que lhe cabe historicamente na administração pública e entre os seus congêneres.

Este caráter da instituição manifesta-se, entre outros aspectos, pela ampliação da gama de serviços prestados na área cultural, dentre os quais destaca-se a publicação sistemática de instrumentos de pesquisa e de manuais técnicos que auxiliem no processo de formação e aperfeiçoamento dos profissionais da área de arquivos.

O Arquivo do Estado tem-se valido, para esse trabalho, de valiosas parcerias, cabendo especial relevo às mantidas com a Imprensa Oficial e a Associação de Arquivistas de São Paulo - ARQ/SP. A primeira delas vem possibilitando o desenvolvimento de um extenso programa de publicações e a segunda tem resultado numa assessoria permanente - tanto formal, quanto informal - na área da Arquivística, eis que a ARQ/SP congrega especialistas de renome internacional nesse campo de atuação.

A presente publicação é, pois, o resultado do esforço conjugado das três instituições e integra o Projeto “Como Fazer”. Orientado, como o próprio nome indica, para aspectos práticos do dia-a-dia dos profissionais da área, esse projeto prevê uma série de outras publicações, sempre de autoria de professores com larga experiência na organização de arquivos.

A direção e o corpo técnico do Arquivo do Estado sentem-se gratificados pelos excelentes frutos já colhidos desse profícuo relacionamento, contando que o mesmo se perpetue e se intensifique, em benefício da comunidade arquivística e da cultura em nosso Estado.

Dr. Fausto Couto Sobrinho  
Diretor do Arquivo do Estado

## SOBRE AS AUTORAS

**Johanna Wilhelmina Smit** é docente de graduação e pós-graduação do Departamento de Biblioteconomia e Documentação da Escola de Comunicações e Artes da Universidade de São Paulo (USP). Formada em Biblioteconomia pela USP, fez Mestrado e Doutorado na França, na área de análise documentária. É co-coordenadora e professora do Curso de Especialização em Organização de Arquivos, promovido pelo Instituto de Estudos Brasileiros e pela Escola de Comunicações e Artes da USP, membro do Conselho Técnico do Sistema de Arquivos da USP (SAUSP) e consultora no Sistema de Arquivos do Estado de São Paulo (SAESP).

**Nair Yumiko Kobashi** é docente de graduação e pós-graduação nos cursos de Biblioteconomia da Escola de Comunicações e Artes da Universidade de São Paulo (ECA/USP) e da Pontifícia Universidade Católica de Campinas (PUC/CAMP). Formada em Biblioteconomia e Jornalismo pela Universidade de São Paulo, fez Mestrado e Doutorado na área de Análise Documentária. É consultora de projetos que envolvem a construção de vocabulários controlados.

## O QUE É CONTROLE DE VOCABULÁRIO

A massa documental que circula por uma instituição ou foi recolhida a um arquivo permanente pode ser comparada às cidades medievais que, construindo muros e/ou fossos, protegiam-se dos perigos externos. Os muros tinham algumas portas e os fossos, algumas pontes, pois, para além da proteção, o acesso à cidade era igualmente necessário.

Os documentos não têm inimigos externos contra os quais devem se proteger, mas sua acumulação acarreta o surgimento de inimigos internos: a falta de organização ou a desorganização do arquivo e a conseqüente dificuldade ou impossibilidade de nele encontrar os documentos procurados. Para que um arquivo, qualquer um, preencha sua missão de ser “o arsenal da administração e o celeiro da história”,<sup>1</sup> pressupõe-se a sua organização. Organizar, segundo o dicionário, significa “dispor de forma ordenada; dar forma regular às partes de um todo; prover de ou adquirir certa estrutura”,<sup>2</sup> ou seja, organizar significa, forçosamente, categorizar, incluir, em função de algum critério, determinados elementos em uma classe e, nesse mesmo movimento, excluir outros elementos.

O conceito arquivístico da série, enquanto “seqüência de unidades de um mesmo tipo documental”,<sup>3</sup> exemplifica o que foi acima afirmado: ao erigir o tipo documental como critério para

---

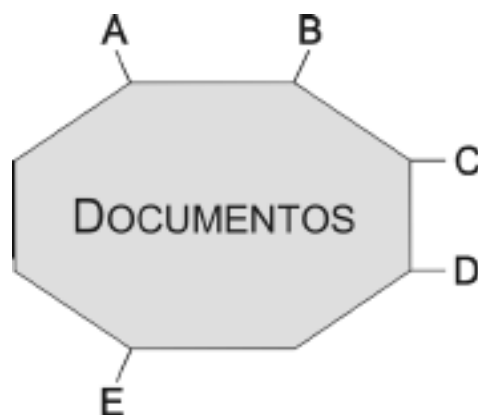
<sup>1</sup> Expressão do arquivista francês Charles Braibant, transcrita por Heloísa Bellotto em *Arquivística: objetos, princípios e rumos*. São Paulo: Associação de Arquivistas de São Paulo, 2002. (Scripta, 1). p. 9.

<sup>2</sup> *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2001.

<sup>3</sup> CAMARGO, A. M. de A., BELLOTTO, H. L. *Dicionário de terminologia arquivística*. São Paulo: Associação dos Arquivistas Brasileiros – Núcleo Regional de São Paulo/Secretaria de Estado da Cultura, 1996.

identificação das séries, fica estabelecido que documentos de um mesmo tipo documental apresentam uma característica comum capaz de diferenciá-los de outros tipos documentais. Assim, agregam-se documentos do mesmo tipo documental numa série, excluindo-se conseqüentemente, e por definição, outros tipos documentais. A organização pressupõe, portanto, a identificação de semelhanças e de diferenças.

Voltando à imagem das cidades medievais evocada no início deste manual, e lembrando que os arquivos e sistemas de arquivos devem organizar um volume importante de documentos, torna-se necessário identificar seus agrupamentos. Eis a função dos “pontos de acesso”, que nada mais são do que as portas nos muros medievais ou as pontes que permitem ultrapassar o fosso. Os pontos de acesso permitem, como a expressão indica, o acesso aos documentos a partir de alguns pontos preestabelecidos, nomeando critérios utilizados para agrupar documentos. Dito em outros termos, os pontos de acesso permitem detectar, no todo da massa documental, documentos correspondentes a determinados agrupamentos.



Os pontos de acesso A, B, C, D e E – na função de portas em muros ou pontes sobre fossos – nomeiam características distintivas dos documentos ou conjuntos documentais, como por exemplo:

- tipo documental;
- órgão de origem e
- data.

O acesso à informação nos arquivos é mediado por pontos de acesso, portas ou pontes que permitem detectar agrupamentos de documentos, distinguindo-os de outros agrupamentos de documentos.

O controle de vocabulário intervém na organização dos arquivos ao nomear, de forma consistente, os pontos de acesso aos documentos e à informação neles contida. O objetivo a ser alcançado pelos arquivos, por essa óptica, é sempre o da recuperação da informação: somente esse objetivo justifica os cuidados com o controle de vocabulário.

Os documentos, ou conjuntos de documentos, devem ser identificados por algumas características que os individualizam, distinguindo-os dos demais: essa é a condição para que possam ser recuperados com precisão. Logo, a identificação das características dos documentos pressupõe:

- que a descrição das características dos documentos seja eficiente, ocupando o menor tempo possível, tanto do arquivista quanto do usuário do sistema;
- que a descrição das características dos documentos seja eficaz, permitindo que o objetivo do arquivo seja atingido, isto é, a correta recuperação dos documentos e informações.

Trata-se, portanto, de identificar, nos documentos ou seus conjuntos, as categorias informacionais que serão utilizadas na busca de informação, ou seja, os pontos de acesso. As normas ISAD (G) e ISAAR (CPF), por exemplo, dispõem categorias informacionais para a descrição de conjuntos documentais ou documentos isolados, respeitando os princípios arquivísticos.

Para tornar úteis os pontos de acesso, duas condições devem ser observadas:

1. que sejam bem selecionados, isto é, que tenham poder discriminante para diferenciar conjuntos documentais em conformidade com os preceitos arquivísticos; essa é a função das normas ISAD (G) e ISAAR (CPF);
2. que sejam expressos de forma consistente.

Assim, para ilustrar as duas condições acima referidas, pode-se dizer que a identificação do órgão gerador/acumulador da documentação constitui um ponto de acesso pertinente. No entanto, para que a busca de informação acerca de determinado órgão gerador de documentos seja feita de forma precisa, gerando uma resposta consistente e, por conseguinte, confiável, é imprescindível que o nome do órgão gerador tenha sido preenchido de forma uniforme. Se os documentos gerados pela Associação de Arquivistas de São Paulo são identificados no campo do órgão gerador, algumas vezes pelo nome por extenso, outras vezes pela sigla “ARQ-São Paulo” ou, ainda, Associação dos Arquivistas de São Paulo, a recuperação da informação será falha, incompleta, gerando resultados equivocados.

Este manual não discutirá os diferentes pontos de acesso, mas as vantagens advindas da adoção do controle de vocabulário no seu preenchimento.

O controle de vocabulário é um recurso para organizar e recuperar documentos – e informações – com consistência, gerando, conseqüentemente, confiança no sistema.

De fato, o aumento no volume de documentos armazenados nas organizações (públicas e privadas) e a sua necessidade cada vez maior de recuperação rápida e precisa tornam patente que a solução passa, forçosamente, por uma classificação rigorosa. A expressão “classificação rigorosa” remete, tradicionalmente, a um plano de classificação que possa ser considerado adequado em relação à instituição, ou seja, um plano que organiza a guarda – física – dos documentos ou processos de tal forma que estes possam sempre ser encontrados quando requisitados.

A organização física dos documentos, embora deva continuar sendo feita de forma rigorosa, não é, de modo geral, suficiente para garantir a recuperação da informação em diferentes situações. O caos instala-se com relativa facilidade, e raramente o vilão é corretamente identificado: o controle de vocabulário – ou o seu descontrole.

Para melhor dimensionar a questão aqui apontada, basta refletir a respeito do uso crescente das tecnologias de informação nas

organizações. A informatização na produção dos documentos, em sua busca, o controle de sua tramitação (protocolo eletrônico) e o acesso aos documentos eletrônicos não podem hoje ser ignorados; afinal, são a realidade de muitas instituições. Nesses casos, a organização física dos documentos pouco significa e em nada colabora para sua recuperação.

Voltemos aos documentos (em suporte papel ou eletrônicos) e sua identificação: se alguns deles são registrados sob a rubrica “solicitação de transporte” outros sob a rubrica “pedido de transporte” e ainda alguns sob a rubrica “requerimento de veículo”, uma busca pelo termo “carro” não produzirá nenhuma resposta, e uma busca pelo termo “transporte” irá gerar uma resposta incompleta, pois outros documentos, que também deveriam ser encontrados, serão, literalmente, ignorados pelo sistema, como se não existissem. O exemplo é simples, e veremos que ele não representa toda a complexidade do controle de vocabulário; no entanto, permite chegar à seguinte afirmação:

O controle de vocabulário é um meio para produzir confiança no sistema de organização e busca de informações arquivísticas.

O controle de vocabulário foi inicialmente desenvolvido pela documentação e introduziu-se na arquivística, especialmente em razão da informatização crescente dos arquivos.

Nos arquivos correntes, diferentes rotinas incluem – ou deveriam incluir – a preocupação com o controle de vocabulário:

- na elaboração de tabelas de temporalidade dos documentos, através da nomeação das séries documentais;
- na elaboração de planos de classificação, através da nomeação das atividades desenvolvidas pela instituição ou, ainda, na nomeação das séries documentais;
- na elaboração de listas de assuntos para protocolos automatizados;
- na atribuição de títulos a formulários.

Nos arquivos correntes o controle de vocabulário deve ser estabelecido em decorrência do diálogo entre o arquivista e os produtores de documentos. O vocabulário deve, nesses casos, aproximar-se tanto quanto possível da linguagem adotada pelos agentes da instituição.

Nos arquivos permanentes, por sua vez, o vocabulário também deve ser estabelecido, mas o arquivista não poderá dialogar com os produtores dos documentos, e o vocabulário se defrontará com questões diacrônicas complexas: a mesma função, ou atividade, com o passar do tempo pode ser nomeada de formas diferentes. Acresce a essa dificuldade uma outra, de mesma natureza: a linguagem utilizada pelo pesquisador, que vem ao arquivo em busca de documentos, também representa uma diferença em relação à linguagem presente nos documentos, tornando a fixação do vocabulário ainda mais difícil. A elaboração de índices (onomásticos, geográficos, de assuntos, etc.) em instrumentos de pesquisa pressupõe igualmente o controle de vocabulário.

Dito em outros termos, a abertura de um processo, a nomeação de um documento ou de uma série documental, a busca de um documento pelo protocolo automatizado ou pelo índice de um instrumento de pesquisa serão operações mais consistentes e confiáveis se forem baseadas num vocabulário controlado.

Para melhor dimensionar a dispersão ocasionada por uma nomeação ou busca de documentos sem a adoção do controle de vocabulário, voltemos ao exemplo citado anteriormente. Imaginemos que os pedidos de carros sejam nomeados de cinco formas diferentes:

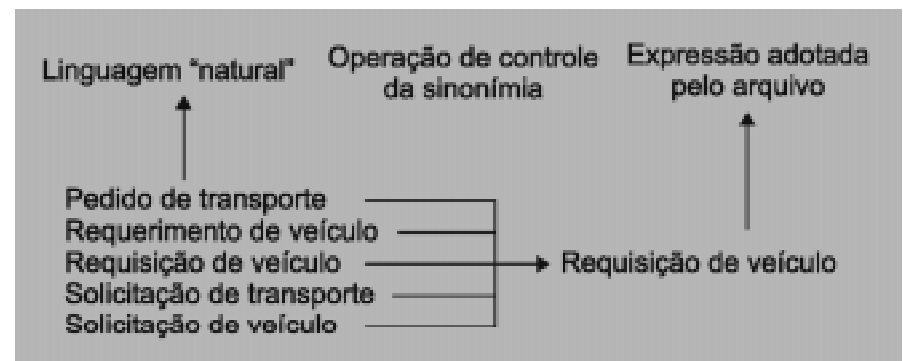
- pedido de transporte;
- requerimento de veículo;
- requisição de veículo;
- solicitação de transporte e
- solicitação de veículo.

Conseqüências? Múltiplas, e nunca agradáveis. Se, por exemplo, ocorrerem problemas no agendamento de um carro, para localizar o respectivo documento e verificar o que aconteceu, será necessário lembrar como o pedido foi nomeado no momento da

solicitação. E quando o Chefe da Seção de Transportes quiser fazer um estudo a fim de avaliar se a frota de veículos está bem dimensionada em face da demanda, ele deverá lembrar que os dados necessários para sua análise se encontram dispersos por cinco séries documentais. Com o objetivo de evitar esses dissabores, a solução mais racional passa pelo controle de vocabulário:

- na entrada do sistema, com a adoção de uma única forma de designação para nomear documentos que são gerados pela mesma atividade;
- na saída do sistema (na busca), informando como cada atividade é nomeada pelo sistema.

Esquemáticamente, a questão pode ser representada como uma operação de redução – e controle – de sinonímias e adoção de uma expressão considerada mais adequada:



Trata-se, em suma, de utilizar um sistema de referência comum e compartilhado (o vocabulário controlado), tanto para organizar quanto para recuperar documentos estocados no arquivo.

Muito embora o controle de vocabulário deva ser incorporado à elaboração dos instrumentos de gestão dos arquivos, este manual, a partir de agora, enfocará especialmente (mas não exclusivamente) a nomeação de funções e de atividades que integram os planos de classificação.



### 1.1. Controle e Uso Controlado de Vocabulário

É muito importante diferenciar claramente:

- o momento da construção de um vocabulário controlado;
- o momento da utilização do vocabulário controlado.

Os dois momentos (construção e utilização de vocabulários controlados) são definidos por diferentes variáveis:

OPERAÇÃO/MOMENTO	VARIÁVEIS
Geração do vocabulário controlado	<ul style="list-style-type: none"> <li>- Política da instituição como um todo (sua missão e objetivos)</li> <li>- Atividades/funções envolvidas na documentação a ser organizada (linguagem de especialidade)</li> </ul>
Utilização do vocabulário controlado	<ul style="list-style-type: none"> <li>- Política de acesso à informação →</li> <li>Política de indexação</li> <li>- Público-alvo</li> </ul>

Enfatizaremos, neste manual, os procedimentos para a construção de vocabulários controlados, lembrando sempre que esses vocabulários são instrumentos que concorrem para, na organização do arquivo, evitar a dispersão de informações e, por consequência, gerar confiança na resposta. Em outros termos, tal construção é um procedimento anterior à sua utilização.<sup>4</sup> Voltemos, pois, a ela, supondo sua utilidade estabelecida, mas lembrando que, idealmente, não deve ser feita por uma única pessoa, mas por uma equipe multidisciplinar.

Como a construção de vocabulário controlado representa uma operação complexa que envolve, de forma ideal, uma equipe de

<sup>4</sup> Uma parcela da bibliografia, no entanto, apresenta a construção de um vocabulário controlado como uma operação concomitante à sua utilização, o que é totalmente contra-indicado, a nosso ver, tendo em vista que as variáveis envolvidas nos dois momentos divergem totalmente.

profissionais, é importante sempre verificar se já existe um vocabulário adequado aos propósitos perseguidos. Porém, tal recomendação esbarra em duas dificuldades:

1. em função da introdução relativamente recente da problemática na área, ainda não se dispõe de um número razoável de vocabulários controlados desenvolvidos para arquivos;
2. as empresas privadas que desenvolvem seus próprios vocabulários controlados, via de regra, não os colocam em domínio público, pois consideram resultado de um investimento.

Com o intuito de contornar as numerosas dificuldades que a questão evoca, sugere-se, para a construção de vocabulários que possam ser adotados pelas diferentes instituições, grupos de trabalho, formados por profissionais de arquivos similares. O intercâmbio entre os arquivos seria fortalecido, gerando vocabulários menos locais (ou seja, adotados por uma única instituição), mais visíveis e com menor custo em função do compartilhamento de profissionais e recursos.

A título de exemplo, pode-se imaginar que os arquivos municipais de determinado Estado construam um vocabulário controlado das atividades-meio desenvolvidas pelas diferentes secretarias municipais. Os arquivos das delegacias de ensino poderiam desenvolver um vocabulário a ser adotado por todos os órgãos a elas subordinados, assim como os arquivos das diferentes secretarias de cultura dos Estados brasileiros, presume-se, gerenciem tipos documentais que se repitam em considerável proporção.

A existência de algumas peculiaridades em cada órgão, na geração de documentos, não invalida a sugestão acima, pois acredita-se que, entre órgãos semelhantes, certamente uma boa porcentagem de funções/atividades desenvolvidas e a geração dos respectivos tipos documentais são constantes e poderiam ser objeto de um vocabulário controlado. As peculiaridades de cada órgão seriam acrescentadas ao vocabulário, em espaços predeterminados e de acordo com os procedimentos acordados pelas diferentes instituições envolvidas.

De toda maneira, a organização de um arquivo supõe a adoção de um vocabulário controlado, fato que requer a compreensão de suas características, quer seja para avaliar vocabulários existentes, quer seja para criar novos ou utilizá-los.

Assim, temos que:

- **CONTROLE DE VOCABULÁRIO** = processo, um objetivo que se deseja atingir;
- **VOCABULÁRIO CONTROLADO** = um instrumento para nomear as atividades/funções, gerando confiança no sistema.

### 1.2. Variáveis intervenientes na construção de um vocabulário controlado para arquivos

Como será a seguir detalhado, a construção de vocabulários controlados mobiliza uma série de procedimentos visando à normalização dos termos (item 2) e à ordenação destes (item 3). No entanto, cumpre preliminarmente passar em revista as variáveis mais freqüentes que nortearão a construção do vocabulário:

- **Tamanho do arquivo e projeção de crescimento, no que diz respeito à variedade de atividades e tipos documentais nele presentes** — de fato, não se trata de tirar conclusões da metragem ocupada pela massa documental, mas da complexidade e diversidade de funções desenvolvidas pela instituição:

Quanto mais complexas e diversificadas forem as instituições, maior será o número de termos do vocabulário controlado necessário para representar adequadamente a variedade do universo documental a ser organizado.

- **Área de atuação da instituição.** Trata-se de analisar se as atividades-fim da instituição são muito especializadas:

No mesmo intuito de bem representar o universo documental, atividades-fim muito especializadas supõem a incorporação de uma linguagem mais especializada.

- **Objetivos da instituição e determinação de prioridades no atendimento por parte do arquivo.** A quem o arquivo deve servir? Deve-se acrescentar que a organização de um arquivo e a respectiva adoção de um vocabulário controlado supõem escolhas em termos do código lingüístico:

O ideal é que o vocabulário controlado reflita a linguagem da instituição, aproximando-se o máximo possível da linguagem do “usuário”.

Se a instituição for muito grande e complexa, a questão se torna mais nuançada, porque os profissionais envolvidos nas rotinas arquivísticas não detêm, forçosamente, o mesmo conhecimento e linguagem especializada do profissional que solicita e utiliza a informação arquivada.

**Todas as variáveis evocam questões de linguagem e sua adequação a culturas locais, ou seja, a composição sociocultural e socioprofissional dos usuários do sistema.**

Em conclusão, a nomeação das séries documentais, embora ancorada em terminologias das áreas de direito, administração e recebendo importante subsídio da diplomática, aponta para uma questão específica que deve ser incluída nas discussões arquivísticas, questão esta até agora referenciada pela expressão “controle de vocabulário” ou pelo instrumento para atingi-lo, isto é, o “vocabulário controlado”.

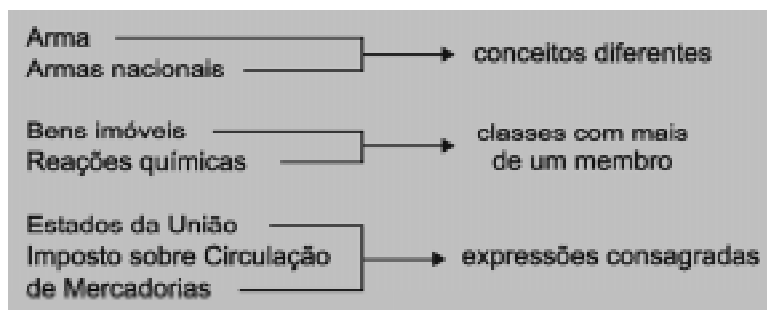
# COMO FAZER O CONTROLE DE VOCABULÁRIO: PROCEDIMENTOS “MICRO”

Os procedimentos “micro” de controle de vocabulário têm por finalidade introduzir o controle nos termos ou expressões e entre estes. Estes procedimentos contemplam, particularmente, seis aspectos:

- Normalização gramatical (item 2.1.);
- Opções de grafia (item 2.2.);
- Alterações em nomes de pessoas ou topônimos (item 2.3.);
- Controle de sinonímia, homonímia e quase-sinonímia (item 2.4.);
- Adoção de termos compostos (item 2.5.) e
- Introdução de notas de escopo e de uso (item 2.6.).

## 2.1. Normalização gramatical

A normalização gramatical preconiza que se adote a forma substantiva, masculina e singular dos termos. Porém, algumas exceções podem ser mencionadas, no que diz respeito à forma plural, quando o termo representar, nessa forma, um conceito diferente ou uma classe de objetos, ou, ainda, quando, em termos compostos, uma parte dos mesmos estiver no plural e a expressão for consagrada:



## 2.2 Opções de grafia

As opções de grafia ocupam um papel de destaque nas discussões arquivísticas relacionadas ao controle de vocabulário. Trata-se, em suma, de resolver situações nas quais o mesmo termo ou expressão apresenta grafias diferentes, freqüentemente em função da passagem do tempo. Assim, por exemplo, o Rio Tamaduateí já foi grafado Tamanduatehy: mesmo rio, mas duas grafias diferentes. Situações assim podem ser resolvidas de duas formas distintas, alertando-se para o fato de que, ao elaborar um vocabulário, dever-se-á manter a mesma opção ao longo de todo o processo.

A primeira forma de encaminhamento consiste em adotar a grafia atual, remetendo das anteriores para esta:

Tamanduatehy, rio **USE** Tamanduateí, rio  
|  
Grafia adotada

Em arquivos permanentes, em que essas questões aparecem com mais freqüência, pode ser interessante adotar uma postura “pedagógica” e, no vocabulário controlado, incluir uma menção a respeito da mudança de grafia:

Tamanduateí, rio - grafia adotada a partir de ..... , em substituição à grafia anterior Tamanduatehy

A segunda alternativa consiste em respeitar as grafias de época, mas relacioná-las, a fim de manter o controle sobre o todo:

Tamanduateí, rio (grafia utilizada em documentos a partir de ... – em documentos anteriores **USE** Tamanduatehy, rio)

Tamaduatehy, rio (utilizado em documentos até ... – em documentos posteriores **USE** Tamanduateí, rio)

A título de sugestão recomenda-se a adoção da primeira alternativa, partindo do pressuposto de que a busca será feita, na

maior parte das vezes, a partir da grafia atualizada do termo. É importante ressaltar que o uso de remissivas alivia o peso das decisões, sendo necessário, ainda, manter a coerência em relação ao critério adotado.

### 2.3. Alterações nos nomes de pessoas ou topônimos

As alterações que ocorrem, nos nomes de pessoas ou topônimos, constituem uma variante das opções de grafia apresentadas no item anterior. No caso de lugares (cidades, bairros, ruas, praças, etc.), supõe-se sempre a possibilidade de resgatar o ato que determinou a mudança (decisão de Câmara, Assembléia, etc.) e, nesse caso, recomenda-se explicitar tal informação:

Floriano Peixoto, rua - até 1907 chamada Fundação, rua da

Fundação, rua da **USE** Floriano Peixoto, rua, nome do logradouro a partir de 1907

O controle de nomes próprios de pessoas instaura uma questão complexa, particularmente no que diz respeito a nomes artísticos ou outras mudanças no nome. O Visconde do Bom Retiro, por exemplo, nasceu José da Silva. Qual nome adotar? Provavelmente o de Visconde, pois supõe-se que, a partir da obtenção do título nobiliárquico, a geração dos documentos tenha sido maior. As normas preconizam a adoção do nome mais conhecido; no entanto, se esta regra pode satisfazer o bom senso, na prática é questionável, pois nem sempre há consenso sobre o nome mais conhecido. A fragilidade da regra aparece claramente quando se trabalha com nomes de pessoas que viveram há muito tempo, em outra região ou país. Por exemplo, como Pedro Álvares Cabral ficou conhecido no Brasil? E em Portugal? E na Espanha?

De toda maneira, em nenhuma hipótese duas regras não podem ser esquecidas:

1. manter a mesma lógica em todos os casos;
2. utilizar o recurso da remissiva para garantir o controle e evitar que a mesma pessoa tenha seus documentos acumulados em dois conjuntos distintos.

Forma adotada

VISCONDE DO BOM RETIRO - nascido José da Silva, tornou-se visconde em .....

JOSÉ DA SILVA **USE** VISCONDE DO BOM RETIRO, título obtido em ....

Remissiva da forma não adotada para a forma adotada pelo sistema

### 2.4. Controle de sinonímia, homonímia e quase-sinonímia

A sinonímia, homonímia e quase-sinonímia constituem um capítulo à parte no controle de vocabulário, respondendo por sua maior visibilidade.

Um exemplo de **sinonímia** anteriormente introduzido foi o de transporte. Como se deve agir nesses casos? Pedido de transporte, requerimento de veículo, requisição de veículo, solicitação de transporte ou solicitação de veículo?

Nesses casos, de sinonímia simples, trata-se, em suma, de fazer uma opção, remetendo da expressão não adotada para a adotada. Não é possível sugerir um critério universal de adoção de um termo ou outro: a adequação da decisão dependerá, prioritariamente, da missão do arquivo, ou seja, da adequação da linguagem adotada em relação a seus usuários. Os casos mais freqüentes de sinonímia que ocorrem em arquivos podem ser sistematizados da seguinte maneira:

- Termos estrangeiros x termos traduzidos  
Feedback ou retroalimentação  
Leasing ou arrendamento mercantil

- Termos mais genéricos x termos mais especializados  
Formatura ou colação de grau

- Siglas x expressão por extenso  
ARQ-SP ou Associação de Arquivistas de São Paulo  
BID ou Banco Interamericano de Desenvolvimento

- Regionalismos

Ar condicionado ou ar refrigerado

Mandioca, aipim ou macaxeira

- Termos adotados pela instituição x termos não adotados pela instituição

Licença-maternidade ou licença-gestante

- Termos adotados pela instituição x termos adotados pela legislação

Licença-saúde ou licença-tratamento de saúde

A variedade de situações acima listada remete tanto à necessidade de conhecimento da cultura local quanto da terminologia adotada pela legislação (último caso). Quando há alguma que rege a atividade cuja nomeação deve ser normalizada, sugere-se a adoção da terminologia sacramentada por essa legislação (e a elaboração de remissivas para variantes). Nos demais casos, somente o conhecimento da cultura da instituição (e, portanto, da linguagem por ela utilizada) permite decidir com adequação. De todo modo, o recurso das remissivas garante o encaminhamento a partir de variantes, evitando assim que as pessoas, na busca, fiquem sem resposta.

A **quase sinonímia** levanta uma questão próxima da sinonímia, porém mais operacional. A rigor, o exemplo supracitado, opondo licença-gestante à licença-maternidade, designa momentos subseqüentes na vida da mulher: primeiro a gestação e depois a maternidade. A menos que os dois momentos fossem totalmente diferenciados pela legislação trabalhista brasileira (caso em que as duas expressões não poderiam ser consideradas sinônimas), operacionalmente pode-se considerar que ambas remetem a uma única atividade: a concessão de licença-gestante.

A **homonímia** (mesma expressão para atividades diferentes) ocorre com freqüência na linguagem do cotidiano, mas é bem menos freqüente na linguagem das instituições, evitando-se desse modo o caos administrativo. No entanto, nas listas onomásticas (de nomes próprios de pessoas), a homonímia ocorre com muita freqüência e deve ser devidamente equacionada. Nesses casos, recorre-se

sempre a especificadores ou modificadores que, através de informações complementares, individualizam os nomes (através do acréscimo do número funcional, número de identidade, ano de nascimento, nome do serviço ao qual estão vinculados etc.).

José da Silva, RG ....., CPF ....., número funcional .....

Qualificativos, entre parênteses, são normalmente empregados para distinguir termos com significados diversos ou evitar interpretações equivocadas, mesmo na ausência do mesmo termo com outro significado:

Ofício (documento)

Ofício (profissão)

Gravador (artista)

Aos topônimos (nomes de lugares) acrescentam-se geralmente informações complementares, mesmo quando não se constata a ocorrência de homônimos. Essa prática é bastante útil, tanto porque explicita tratar-se de nomes de lugares (ou geográficos) quanto por que fornece informações úteis, inclusive para a diferenciação de homônimos.

Direita, rua (Centro)

Direita, rua (Jardim Primavera)

São Paulo (cidade)

São Paulo (Estado)

Triunfo (PE)

Triunfo (RS)

## 2.5. Adoção de termos compostos

Os termos compostos, que representam atividades identificadas por expressões formadas por mais de uma palavra,

geram uma série de dúvidas relacionadas à manutenção, ou não, da expressão como um todo, quanto à ordem a ser adotada entre os termos.

A expressão “administração de arquivos”, por exemplo, suscita duas perguntas:

1. A expressão deve ser considerada indivisível ou deve ser subdividida em duas partes: “administração” + “arquivo”?
2. Caso se mantenha a expressão “administração de arquivos”, é melhor mantê-la na ordem direta (“administração de arquivos”) ou na ordem invertida (“arquivos, administração”)?

A norma ISO 2788 determina que os termos de indexação devem representar o máximo possível de noções simples ou unitárias, e que os termos compostos devem decompor-se em elementos mais simples, salvo se esse procedimento afetar a compreensão do usuário.

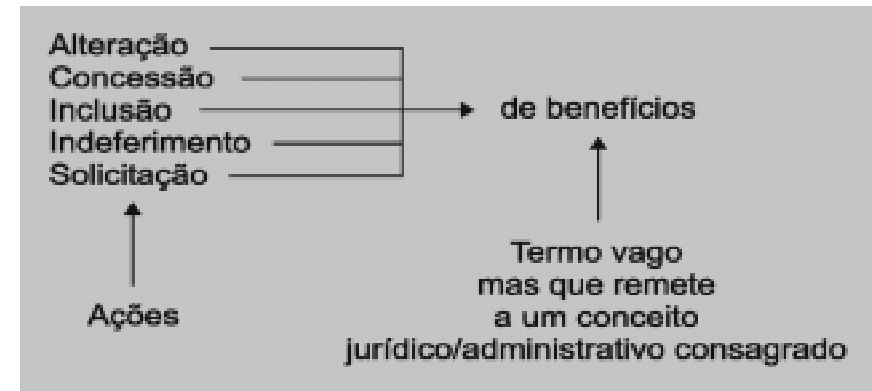
Termos compostos, ou até expressões que se assemelham a frases, são freqüentemente utilizados para nomear tipos documentais ou atividades desenvolvidas pelas instituições. A título de exemplo, basta lembrar as denominações dadas a processos da área pública (“Processo para pagamento de pensão – Revolução Constitucionalista de 1932”) ou a atividades tais como prevenção de acidentes de trabalho ou elaboração de alterações orçamentárias.

Quando se trata de expressões consolidadas, inclusive pela legislação, sugerimos a manutenção da expressão indivisa, pois o todo adquiriu um sentido preciso a ser preservado. Deve-se acrescentar a esse argumento o fato de que, freqüentemente, uma parte da expressão é formada por termos que, isoladamente, significam muito pouco por serem demasiadamente vagos:

- [Licença ?] gestante
- [Licença ?] paternidade
- [Licença ?] sem [vencimentos?]

No entanto, nem todas as expressões são consolidadas pela legislação, e a orientação geral, segundo a qual se deve decidir de forma a não prejudicar a compreensão do usuário, é excessivamente vaga.

Deve-se atentar ainda para outra faceta relacionada aos termos compostos que, quando esses designam atividades, são freqüentemente antecidos por nomes de ações que não incidem no conceito central, mas que se justapõem a ele.



Muito embora a questão dos termos compostos seja seguramente aquela que mais gera dúvidas, como nos casos acima exemplificados, sugere-se a decomposição das expressões. Isso porque os termos que designam ações não significam nada isoladamente. Alteração do quê? Inclusão do quê e onde? Concessão do quê? E assim por diante. A elaboração de uma lista controlada de ações parece factível e útil, pois ajudará a decidir se determinada expressão deve, ou não, ser decomposta, porém ela não deve ser muito extensa.

Tendo em vista a complexidade das decisões a serem tomadas, propomos a adoção das seguintes regras gerais:

1. aproximar-se, tanto quanto possível, da linguagem consolidada pela legislação e da linguagem utilizada pela instituição;
2. valorizar a coerência interna do vocabulário (ou seja, manter a mesma decisão ao longo de todo o vocabulário);
3. diferenciar a nomeação de atividades das ações que as modificam.

Reiterando que a adoção dos termos compostos *com coerência* aponta seguramente para uma das questões mais complexas a enfrentar no controle de vocabulário para arquivos – demandando estudos aprofundados –, chamamos a atenção para a importância da elaboração de listas alfabéticas permutadas, conforme descrito no item 3.4.

## 2.6. Introdução de notas de escopo e notas de uso

No intuito de garantir ao máximo o funcionamento do vocabulário controlado como uma lista monossêmica<sup>5</sup> de termos cujo uso, conseqüentemente, possa se fazer de forma bastante consistente, tanto na indexação dos documentos por ocasião da entrada no sistema quanto no momento da busca (a “saída” do sistema), há dois recursos suplementares que podem ser de grande valia: as notas de escopo e as notas de uso.

O controle de vocabulário pressupõe, efetivamente, além do controle de formas (gramaticais, grafia e termos compostos) e do controle de variantes (sinônimos, homônimos), o controle do significado atribuído a cada termo. Essa nova faceta da discussão parte do pressuposto de que termos designando atividades, mesmo quando estas foram categorizadas, nem sempre garantem uma nomeação monossêmica por parte dos profissionais que nomeiam ou buscam documentos. Quanto maior a instituição, e mais diversificadas as atividades e os profissionais envolvidos, maior a probabilidade de um mesmo termo, que designa uma atividade, gerar entendimentos diferentes. Nesses casos, será necessário conceituar o termo, explicitando a compreensão da atividade nomeada por ele no âmbito da instituição. Vejamos o seguinte exemplo:

<sup>5</sup> A palavra monossêmica tem um único sentido, em oposição às palavras polissêmicas, que têm vários sentidos – DUBOIS, J. et al. *Dicionário de lingüística*. São Paulo: Cultrix, 1991. p. 418.

### ACIDENTES

descritor de atividade

As atividades envolvidas em como lidar com desastres que causam lesão ou dano. Inclui dano ou lesão às propriedades da instituição ou a algum membro da equipe, ocorrido enquanto vinha ou saía do trabalho, ou ao público em geral ou ainda a visitantes nas propriedades da instituição. Também inclui medidas de prevenção de acidentes.

Ver INDENIZAÇÃO - CASOS para situações em que as lesões pessoais resultam em pedido de indenização, e o caso gerar um processo por razões legais ou operacionais. Copiar as informações sobre o acidente e colocá-las no arquivo de indenização. Adicionar o nome do indivíduo em texto livre. Se a pessoa não for um empregado da instituição, adicionar o descritor “visitantes”, e então seu nome em texto livre.

Ver SEGUROS para os processos de retirada dos prêmios de seguro.

Ver RECLAMAÇÕES para gerenciar o dano ou a destruição da propriedade causado(a) por acidentes.

Fonte: National Archives Of Australia. The keyword AAA: a thesaurus of general terms.

Estas explicações têm uma única finalidade: determinar, de forma clara, explícita, o âmbito ou as condições em que determinado termo deve ser utilizado. Não se trata, por extensão, de concordar ou de discordar da leitura da atividade “acidentes” proposta pelo Arquivo Nacional da Austrália, mas de incorporar à compreensão do vocabulário controlado uma nova faceta: a conceituação. Essa conceituação, tal como consta no exemplo acima, constitui o objeto das “notas de escopo”.

Alguns vocabulários distinguem de forma explícita dois tipos de explicações:

1. as **notas de escopo** têm por finalidade explicitar a amplitude ou o entendimento atribuído ao conceito. Elas não podem ser

confundidas com definições de dicionário, pois não se pode perder de vista que os vocabulários controlados têm objetivos concretos e “industriais”: como a organização do acesso à informação contida nos documentos. Desse modo, as definições contidas nas notas de escopo têm objetivos pragmáticos (“organizar o acesso à informação”) e não teóricos;

2. as **notas de uso**, pouco utilizadas na prática, não se propõem a explicitar o conceito mas a explicitar recomendações práticas que devem nortear o uso do termo. No exemplo supracitado, os parágrafos que iniciam por “VER” caracterizam notas de uso.<sup>6</sup>

Um outro exemplo, retirado do plano de classificação em fase de revisão das atividades da USP, talvez ajude a distinguir e a dimensionar a utilidade dos dois tipos de notas explicativas:

#### **Aluno-monitor de graduação**

*Nota de escopo:* As funções de monitor poderão ser exercidas por alunos matriculados em curso de graduação que tenham obtido bom rendimento em disciplinas já cursadas.

*Nota de uso:* utilizar o termo para alunos de graduação, exclusivamente.<sup>7</sup>

O exemplo apresentado indica que a nota de escopo e a nota de uso são complementares e não excludentes.

---

<sup>6</sup> Particularmente, não sugerimos a adoção da expressão VER para introduzir notas de uso, pois esta é usada por outros vocabulários a fim de representar uma relação de sinonímia, conforme sistematizado no anexo.

<sup>7</sup> A USP prevê a figura do aluno-monitor tanto para alunos de graduação quanto para alunos de pós-graduação. O vocabulário, para evitar a homonímia, criou os termos “aluno-monitor de graduação” e “aluno-monitor de pós-graduação”.

#### **2.7. Síntese dos procedimentos “micro” de controle de vocabulário**

O controle de vocabulário que, como o nome já diz, visa estabelecer um vocabulário controlado para a identificação e subsequente busca de documentos e informações no arquivo, com precisão e confiabilidade, processa-se principalmente através dos seguintes procedimentos:

- distinção entre termos adotados e termos não adotados pelo sistema;
- remissivas dos termos não adotados para os termos adotados pelo sistema;
- padronização formal dos termos adotados;
- conceituação de termos.



## COMO FAZER O CONTROLE DE VOCABULÁRIO: PROCEDIMENTOS “MACRO” — ORDENAÇÃO DOS TERMOS

A ordenação dos termos introduz no controle de vocabulário um outro aspecto igualmente indispensável, pois, através dele, os termos são organizados num sistema signficante. Caso não se proceda à ordenação dos termos, o vocabulário controlado se apresentará como uma longa lista de termos, controlados, por certo, mas cujo significado se restringe aos próprios termos. Não é possível extrair significado de uma lista, gerar expectativas ou previsões sobre o universo abrangido, o ponto de vista adotado e nem o nível de especificidade no qual a documentação foi tratada. Assim como uma lista telefônica de determinada cidade não permite fazer inferências sobre as características da população, uma lista de termos controlados não possibilita uma visão do universo compreendido pela lista. Recomenda-se, portanto, que os termos, uma vez submetidos ao controle de vocabulário, sejam ordenados, organizados ou categorizados. A categorização gera significado ao introduzir os termos num sistema signficante.

Como a inclusão de termos em hierarquias é freqüentemente operada, inclusive no senso comum, tem-se tendência a esquecer seu principal produto, ou seja, uma operação de significação.

Isto é, a ordenação, hierarquização, estruturação ou categorização de termos:

- gera a arquitetura do universo de funções e atividades da instituição;
- atribui sentido ao todo.

A organização do universo documental pode ser representada por uma diversidade de vocabulários controlados, como será visto nos itens seguintes.

### 3.1. Tipologia de vocabulários controlados

Há vários tipos de vocabulários controlados, e é importante identificar o mais adequado para cada situação. Resumindo, a tipologia se estabelece em função do princípio de organização dado aos termos do vocabulário, desde a organização por ordem alfabética (a rigor, não se trata neste caso de uma organização), passando por sistemas mais rígidos (planos de classificação) até chegar, historicamente, a sistemas com uma organização menos draconiana (os tesouros). Há possibilidades intermediárias, como listas de atividades organizadas em grandes categorias, em que as atividades são dispostas em ordem alfabética; porém, na discussão que segue, ressaltaremos as características, virtudes e desvantagens de planos de classificação e tesouros, tendo em vista o fato de as discussões enfocarem com mais freqüência esses dois tipos de vocabulários controlados.

Iniciemos pela análise de uma situação concreta, retirada do cotidiano de uma metrópole, e enfocando o deslocamento nessa cidade entre dois pontos predeterminados (a residência e o local de trabalho, por exemplo). Esse deslocamento entre dois pontos pode ser focado a partir de dois paradigmas: os roteiros das linhas regulares de ônibus e o percurso realizado com uma bicicleta ou um carro:

Paradigma	Características	Conclusão
Roteiro de linhas regulares de ônibus (transporte coletivo)	Menos opções de trajetos Nenhuma adaptabilidade (os trajetos estão previamente determinados) Os itinerários e as paradas foram previstos <i>a priori</i>	Sistema menos flexível, pois todas as variáveis (trajetos e paradas) devem ser previstas <i>a priori</i>
Percurso com carro ou bicicleta (transporte individual)	Mais opções de trajetos Adaptabilidade no caso de imprevistos Itinerário variável e, portanto, menos previsível	Sistema mais flexível, pois as combinatórias não precisam ser previstas <i>a priori</i> mas podem ser estabelecidas de acordo com a necessidade

O paradigma do roteiro de ônibus representa bem o conceito de um sistema classificatório, enquanto o percurso com carro ou bicicleta simboliza o tesouro. Não há razão para interpretar esses dois tipos de vocabulários controlados como opostos radicais, mas interessa verificar melhor o que os distingue.

O **plano de classificação** tem por base o princípio da hierarquia, que se aplica a todas as atividades desenvolvidas pela instituição. Esse princípio da hierarquia tem uma grande vantagem: ao ordenar as atividades hierarquicamente, ele fornece, instantaneamente, uma visão do todo e também de como este foi distribuído. No entanto, a vantagem pode se tornar uma desvantagem quando se trabalha para ordenar um universo maior e, por definição, mais complexo. Simplificando o princípio do sistema classificatório, parte-se do pressuposto de que todas as atividades devem ser inseridas, ou alocadas, em classes, sendo que essas classes devem ser auto-excludentes. Não se evoca, aqui, a diferença entre o método estrutural e funcional para classificar os documentos, pois ambos devem se conformar aos mesmos pressupostos da teoria da classificação.

O exemplo do percurso com carro ou bicicleta simboliza, por sua vez, a lógica do tesouro, que não abandona a lógica do plano de classificação, apenas a abranda ou a torna mais relativa. O todo (conjunto das funções/atividades desenvolvidas) é segmentado em conjuntos menores (as categorias) e, no interior destas, organizam-se os termos, explicitando-se as relações entre eles (hierarquia, equivalência, associação espaço-temporal), o que torna a sua atualização mais simples e o seu uso, mais consistente. Obtêm-se, desse modo, mapas parciais, que permitem vários trajetos, sendo, portanto, menos coercitivos ou mais flexíveis em sua aplicação. Um exemplo é da área de arquivos universitários, comparando-se a solução dada para a rubrica “bolsas de estudo” nos planos de classificação da Universidade Carlos III (Madri, Espanha) e da Universidade de São Paulo.<sup>8</sup> Com a solução proposta pelo tesouro da Case Western Reserve University (Estados Unidos) para a mesma rubrica, fica mais clara essa diferença.

UNIVERSIDADE CARLOS III DE MADRID Cuadro de Clasificación de los Documentos	
<b>L100 GESTIÓN DE LA INVESTIGACIÓN</b>	
<b>L101 Organización y programación de la investigación</b>	
L102 Proyectos de investigación	
L103 Programación de la investigación	
<b>L104 Ayudas a la investigación</b>	
L105 Información y difusión de convocatorias	
L106 Convenios y transferencia de tecnología	
L107 Convenios con empresas	
L108 Convenios con centros de investigación	
L109 Contratos	
L110 Contratos de estudios jurídicos	
L111 Solicitud de ayudas	
L112 Infraestructura científica	
L113 Proyectos de investigación	
L114 Becas ←	
L115 Ayudas a la movilidad del personal investigador	
L116 Plan de fomento a la investigación	
L117 Organización y asistencia a congresos	
<b>L118 Difusión de los resultados de la investigación</b>	
<b>L119 Evaluación de la investigación</b>	
<b>F100 GESTIÓN DE LOS RECURSOS ECONÓMICOS</b>	
(...)	
<b>F124 Gestión de los gastos</b>	
(...)	
F128 Gestión contable de la investigación	

<sup>8</sup> O Plano de Classificação das Atividades da USP está sendo revisto e ampliado neste momento. Os exemplos, no entanto, foram extraídos da versão datada de 1997.

UNIVERSIDADE DE SÃO PAULO Plano de Classificação das Atividades	
Q00000	PESQUISA
Q10000	BOLSA DE PESQUISA ←
Q20000	INTERCÂMBIO CIENTÍFICO
Q30000	PÓS-DOCTORAMENTO
Q40000	PROGRAMA DE PESQUISA
Q40200	Divulgação de resultados
Q40300	Linhas de pesquisa
Q40500	Recursos
Q60000	PROJETOS DE PESQUISA
Q60200	Divulgação de resultados
Q60500	Recursos

CASE WESTERN RESERVE UNIVERSITY ARCHIVES Thesaurus of University Terms	
C00000	ACADEMICS
(...)	
C73000	Research
C73380	Human subjects
C73730	Technology transfer
J00000	FINANCES
(...)	
J37000	Grants ←

O plano de classificação da Universidade Carlos III inclui as bolsas de estudo (“becas”) na categoria das “solicitações de apoio”;<sup>9</sup> por sua vez incluem uma subdivisão da categoria “apoio à pesquisa”, já uma subdivisão da grande função “gestão da pesquisa”. O Plano de Classificação da USP, ao subordinar as bolsas de pesquisa à função

<sup>9</sup> O termo apoio é adotado no Brasil para designar ajuda financeira à pesquisa: a tradução não foi, portanto, literal, mas adaptada à cultura local.

“pesquisa”, representa a adoção de um ponto de vista muito semelhante. As bolsas de estudos têm, no entanto, uma dimensão financeira (trata-se de um apoio financeiro que, em contrapartida, favorece a realização de determinada pesquisa): esse aspecto foi silenciado pelos planos de classificação das universidades Carlos III e USP.

O tesouro da universidade americana incluiu as bolsas de estudos (*grants*) na função “finanças”, optando assim por uma abordagem diversa. Esse fato, em si, somente demonstra que as três universidades têm culturas diferentes e, portanto, atribuem um significado diferente às bolsas de estudo. No entanto, a Case Western não ignora que as bolsas de estudo mantêm uma interface com a gestão da pesquisa: o tesouro, na lista alfabética, sob a rubrica “pesquisa” (*research*), aponta para um termo relacionado (RT)<sup>10</sup> *grants*. Ou seja, as bolsas de pesquisa, muito embora categorizadas em função de seu aspecto financeiro, aparecem explicitamente relacionadas, na estrutura do tesouro, à pesquisa:

CASE WESTERN RESERVE UNIVERSITY ARCHIVES Thesaurus of University Terms - lista alfabética	
GRANTS	
BT Finances ←	
RT Research	
(...)	
RESEARCH	
BT Academics	
NT Human subjects	
Technology Transfer	
RT Grants	

O exemplo das “bolsas de pesquisa” permite chegar a duas conclusões:

1. uma categorização hierárquica de termos representa – ou atualiza – sempre um ponto de vista;<sup>11</sup>

<sup>10</sup> Em anexo, encontram-se sistematizados e explicitados os operadores utilizados em planos de classificação e tesouros.

<sup>11</sup> Supõe-se que este ponto de vista mantenha um vínculo muito claro com a cultura institucional.

2. as opções de categorização ao mencionar termos não incluídos em determinada função/atividade, mas considerados relacionados, podem ser explicitadas na lógica do tesouro.

Planos de classificação e tesouros oferecem duas formas – distintas mas não opostas – de apresentação dos termos, dispondo os tesouros de mais recursos para explicitar a modalidade de relação estabelecida entre os termos, conforme está sintetizado no anexo.

O exemplo a seguir, hipotético, permite comparar as duas apresentações:

Plano de Classificação	Tesouro
<b>EVENTOS</b> <b>Eventos científicos</b> Congressos <b>Eventos educativos</b> Conferências Palestras	<b>CONFERÊNCIAS</b> NE use o descritor para reuniões programadas para dar a palavra a uma pessoa de reconhecida competência TG Eventos educativos TR Congressos UP Seminários Simpósios Workshops
	<b>CONGRESSOS</b> NE use o descritor para reuniões programadas para favorecer o intercâmbio científico/acadêmico TG Eventos científicos TR Conferências
	<b>EVENTOS CIENTÍFICOS</b> TG Eventos TE Congressos
	<b>EVENTOS EDUCATIVOS</b> TG Eventos TE Conferências Palestras
	<b>PALESTRAS</b> TG Eventos educativos
	Seminários <b>USE</b> Conferências Simpósios <b>USE</b> Conferências Workshops <b>USE</b> Conferências

Um vocabulário controlado (plano de classificação ou tesouro) é composto, idealmente, de duas partes:

- **uma parte categorizada** (ou estruturada), em que as atividades são ordenadas pelas funções (ou pela estrutura, caso se adote o critério estrutural). Essa parte pode, ainda, incluir as notas que restringem ou explicitam o significado dos termos;

- **uma lista alfabética**, em que as denominações adotadas para as atividades remetem para a lista categorizada (funcionando, portanto como índice), ao passo que o controle de vocabulário se manifesta pela inclusão de termos não adotados (remetendo aos adotados).

A relação que se estabelece entre a parte categorizada e a lista alfabética pode ser visualizada no próximo exemplo, também retirado do universo dos arquivos universitários:

SISTEMA DE ARQUIVOS DA USP –SAUSP	
Plano de classificação	
C00000 ENSINO: GRADUAÇÃO	Lista categorizada: inclui a atividade da "Colação de grau" na categoria maior do "Ensino de Graduação"
: C35000 Colação de grau	
A	Lista alfabética: inclui a atividade da "Colação de grau" na letra C da lista, remetendo à localização na lista categorizada (C35000). Na letra F aparece a menção a um sinônimo "formatura", remetendo deste para o termo adotado, "colação de grau".
B	
C Colação de grau - C35000	
D	
E	
F Formatura <b>USE</b> Colação de grau	

Ao final, o que concluir? De imediato, deve-se atentar para os seguintes pontos:

- um vocabulário controlado é, idealmente, composto de uma lista categorizada e de outra alfabética (que contempla as questões das remissivas);

- a diferença entre planos de classificação e tesouros se manifesta:
- pelo nível de hierarquização: quanto mais o universo das funções/atividades é hierarquizado, mais nos aproximamos do paradigma do plano de classificação;
- pelo fato de que a lista alfabética de tesouros inclui as remissivas, ao passo que a lista alfabética de planos de classificação, quando existente, não manifesta forçosamente essa preocupação.

**RESUMINDO:** um vocabulário controlado, na óptica arquivística, deve:

**1. hierarquizar as funções e respectivas atividades:** esta hierarquia pode ser mais ou menos detalhada, mas é necessária por duas razões:

- para dar uma idéia do universo funcional abrangido (a visão do todo), incluindo os termos num sistema signficante;
- para orientar a organização física das séries documentais.

**2. contemplar o controle de vocabulário, particularmente no que diz respeito à sinonímia,** pela inclusão de remissivas na lista alfabética. As remissivas têm por função orientar a nomeação dos documentos (em sua geração e no momento de sua busca) de termos ou expressões não adotados para termos ou expressões adotados pelo sistema;

**3. relacionar termos** presentes em categorias diferentes.

**4. conceituar os termos e as condições de seu uso.**

**HIERARQUIA CONSISTENTE + SINONÍMIA CONSISTENTE =  
CONFIABILIDADE NA ORGANIZAÇÃO DO ARQUIVO  
CONFIABILIDADE NA RESPOSTA FORNECIDA PELO ARQUIVO**

Na prática dos arquivos, elaboram-se, atualmente, planos de classificação, e somente alguns sistemas, defensores de um acesso mais ágil à informação, preconizam a construção de tesouros. O Arquivo Nacional da Austrália exemplifica, com muita consistência, a segunda opção.

No entanto, de imediato, propomos que se construa um plano de classificação que incorpore a preocupação com o controle de vocabulário e, num segundo momento, seja gerado um tesouro a partir do plano.

Os procedimentos “macro” atribuem significado às atividades desenvolvidas pela instituição através da ordenação do universo. O exemplo das bolsas de estudo, ora incluídas no universo da pesquisa (Universidade Carlos III e USP), ora incluídas nas finanças (Case Western Reserve University), ilustra a atribuição de sentido produzida pela operação da ordenação.

### 3.2. A construção de um plano de classificação

A construção de um plano de classificação que incorpore a preocupação com o controle de vocabulário obedecerá aos seguintes procedimentos:

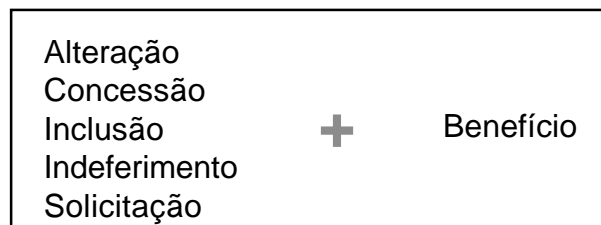
1. levantamento de uma lista livre de atividades. Essa listagem pode ser elaborada a partir da leitura das tabelas de temporalidade dos documentos, da lista de “assuntos” do protocolo, da leitura de nomes dados a processos ou pasta, etc.
2. análise crítica dos termos incluídos na lista: até que ponto a lista duplica atividades em função de sinonímias? Até que ponto tal lista é consistente em termos de normalização gramatical, opções de grafia e solução dada aos termos compostos? Há, por outro lado, sinonímias não contempladas pela lista, mas que se justificariam tendo em vista os hábitos de nomeação de uma parte dos profissionais da instituição?
3. elaboração de uma lista alfabética consistente de atividades, desdobrada nas ações, caso pertinente, e acrescida das remissivas que se fizerem necessárias;
4. categorização, em maior ou menor grau, das atividades constantes da lista. Nomeação das categorias maiores, novamente incorporando a preocupação com o controle de vocabulário;
5. análise das nomeações que podem gerar leituras diferentes e elaboração de notas de escopo ou notas de uso;

6. inclusão de uma codificação (alfanumérica ou simplesmente numérica, mas que permita futuras inserções de novas atividades ou agrupamentos);

7. submissão do vocabulário controlado (lista categorizada e alfabética) a testes, avaliação do resultado dos testes, incorporação de ajustes e efetiva implantação do vocabulário controlado.

### 3.3. A construção de um tesauro funcional<sup>12</sup>

Um tesauro funcional controla o vocabulário que designa a razão de ser do documento (sua função) e não o que ele contém (assunto). A sua construção partirá do plano de classificação e o reordenará. Nunca é demais salientar que um tesauro funcional não substitui tal plano. Na realidade apresenta as diferentes visualizações de um (sua estrutura, a apresentação hierárquica e a lista alfabética). As ações mencionadas no item 2.5, por exemplo, que deviam ser previstas e explicitadas para cada atividade, constituirão uma lista à parte, autônoma: o tesauro não supõe que todas as ramificações da atividade sejam previstas a priori, mas que estas possam ser livremente acopladas às atividades, quando a documentação assim o exigir. Veja-se o exemplo das ações relacionadas ao controle de benefícios (alteração, concessão, inclusão, indeferimento e solicitação), apresentadas no item 2.5. No âmbito de um tesauro, haveria uma atividade “gestão de benefícios” e uma lista, autônoma, de ações – livremente combináveis –, não existindo necessidade de prever, inicialmente, todas as ações que podem ser justapostas à atividade:



<sup>12</sup> Adotamos a terminologia proposta pelo Arquivo Nacional da Austrália.

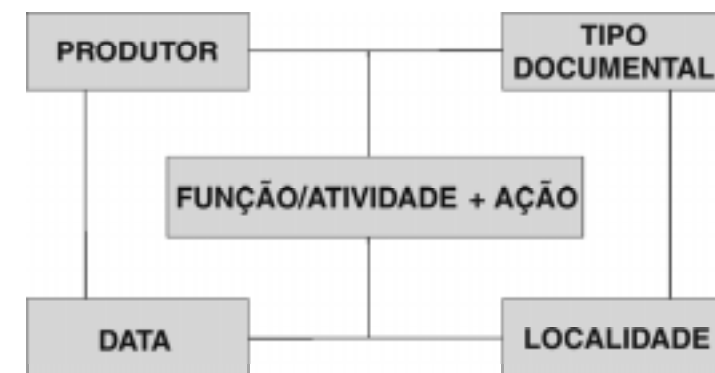
O tesauro, nessa óptica, enfatiza a combinatória entre os termos e, conseqüentemente, não prevê todas as combinações por antecipação, já que as permite quando necessário. A economia do sistema, sua flexibilidade e facilidade de atualização ganham, mas, ao mesmo tempo, perde-se a visão do arquivo como organização de documentos. Em suma, passa-se a trabalhar com dois paradigmas, complementares entre si:

- a organização física dos documentos;
- a organização intelectual das atividades que os produzem.

As funções desempenhadas pelo tesauro funcional podem ser resumidas como segue:

- opera como índice de termos do plano de classificação;
- opera como índice para encontrar séries documentais a serem eliminadas;
- prevê pontos de acesso ao usuário, permitindo navegar de sinônimos ou outros termos não adotados para a terminologia adotada pelo arquivo;
- pode ser usado como ferramenta para ajudar o funcionário a encontrar documentos necessários para as tarefas do dia-a-dia.

O tesauro, ao enfatizar a combinatória entre os termos, apresenta o universo organizado sob a forma de rede, em que cada nó representa um ponto de acesso. Apresentamos a seguir um exemplo de rede:



### 3.4. A elaboração de listas alfabéticas complementares à ordenação dos termos

O plano de classificação e o tesouro, enquanto diferentes formas de apresentação lógica do universo das atividades e dos documentos a serem organizados no arquivo, constituem instrumentos de gestão imprescindíveis. No entanto, pela sua própria apresentação, eles não favorecem a consulta de acordo com a diversidade de linguagens mobilizadas pelos usuários do arquivo para denominar suas necessidades de informação. Dito de outro modo, o plano de classificação e, de forma menos evidente, o tesouro estão voltados prioritariamente para a organização dos documentos no arquivo e não têm por objetivo a função precípua do arquivo, qual seja, a organização do acesso à informação e aos documentos.

Retomando conceitos anteriormente introduzidos, o acesso à informação se dá de diferentes maneiras, de acordo com a origem de quem procura, com o momento em que essa necessidade ocorre e com o objetivo que gerou a busca pela informação. Assim, o administrador, ao final de um ano de gestão, para fazer seu relatório anual, solicitará determinados documentos diferentemente do modo como um outro cidadão buscará informação que diz respeito a um parente que foi funcionário público. O pesquisador, ao enunciar uma necessidade de informação guiada pelo seu objeto de pesquisa, poderá estar procurando pelos mesmos documentos que, anos atrás, foram solicitados pelo administrador ou pelo cidadão, mas o fará certamente de outra maneira, utilizando outros termos. O acesso à informação, portanto, deve incorporar as diferentes maneiras de enunciar necessidades de informação em função da diversidade de públicos, momentos e objetivos aos quais serve.

A lista alfabética de termos, que remete à lista categorizada, dispõe:

- os **termos adotados** pelo sistema em ordem alfabética, ou seja, eles estão acessíveis mesmo para aquele usuário que não sabe como as atividades foram categorizadas ou qual critério foi adotado na categorização (estrutural ou funcional);
- os **termos não adotados** pelo sistema, incluídos na mesma ordem alfabética e encaminhando o usuário para o termo adotado, evitando, desse modo, que fique perdido no labirinto

da organização da informação ou ainda que conclua que a informação por ele procurada não está presente, pelo simples fato de ter recorrido a sinônimos no momento da busca.

A lista alfabética, incluindo termos adotados e não adotados, contempla, dessa maneira, as diferentes formas utilizadas pelos usuários para nomear suas necessidades de informação.

Constatada a importância da geração de listas alfabéticas como acesso à parte categorizada do plano de classificação ou do tesouro, impõe-se uma última reflexão, gerada pela adoção de expressões (ou termos compostos) para nomear atividades. De fato, e retomando as indagações presentes no item 2.5., a nomeação de atividades por termos compostos ou expressões suscita algumas dúvidas suplementares, tão complexas quanto a discussão inserida no item referido. Exemplificando: a inclusão da expressão “licença-tratamento de saúde” na letra L da lista alfabética é correta, mas não prevê uma provável busca por “tratamento de saúde”, na letra T. Os termos compostos também geram, portanto, dúvidas no momento da elaboração das listas alfabéticas, recomendando-se adotar a rotação dos termos, isto é, a segmentação das expressões em partes significativas, e a introdução dessas partes na ordem alfabética:

A  
B  
(...)  
L Licença-tratamento de saúde  
(...)  
T Tratamento de saúde, licença

Ordem alfabética

Dito de outro modo, a expressão é segmentada em partes menores, cada qual incluída na ordem alfabética e complementada pela parte que a antecede:

Adicional por Tempo de serviço

A expressão estará presente, na ordem alfabética, em 2 pontos diferentes:

Adicional por tempo de serviço  
Tempo de serviço, adicional por

As dúvidas ocorrem no momento da segmentação das expressões, pois, se por um lado tal segmentação é necessária, por outro lado é frequentemente difícil decidir quais partes são mais significativas. Assim na expressão “execução de serviços de gráfica”, por exemplo, cabe a dúvida:

Execução de	Serviços de gráfica
-------------	---------------------

ou

Execução de	Serviços de	gráfica
-------------	-------------	---------

Nesses casos, uma vez que tal resolução depende em boa parte da cultura institucional e dos modos de nomeação – variáveis de acordo com os usuários –, ocorre uma única certeza: é melhor pecar por excesso do que por falta. No entanto, a adoção desta regra em demasia torna a lista alfabética muito pesada, ou longa.

As listas alfabéticas construídas dessa maneira são chamadas de índices permutados: “gira-se” a expressão para que cada parte significativa se torne acessível pela ordem alfabética.

## CONCLUSÃO

Para finalizar, sugerimos:

Iniciar pela construção de um vocabulário controlado, incorporando-o ao plano de classificação, pensando no arranjo físico e, num segundo momento, elaborar um tesouro baseado no primeiro vocabulário, priorizando a organização intelectual das funções/ atividades geradoras de documentos.



## O QUE HÁ PARA LER

A bibliografia sobre questões relacionadas ao controle de vocabulário para a área da biblioteconomia e documentação é vasta, mas a literatura que descreve a aplicação do mesmo controle à arquivologia é bastante escassa. Consultamos os seguintes textos:

CALDERON, W. R. *Instrumentos de pesquisa nos arquivos públicos permanentes: um estudo sob a ótica da análise documentária*. São Paulo, 2003. Dissertação (mestrado) - Escola de Comunicações e Artes, Universidade de São Paulo.

RIBEIRO, F. *Indexação e controlo de autoridade em arquivos*. Porto: Câmara Municipal, 1996.

SANTOS CANALEJO, E. C. La indización en la recuperación de la información. *Janus*, n.1, p.116-131, 1998.

Diferentes textos elaborados pelo Arquivo Nacional da Austrália revelaram-se seminiais. Em particular recomendamos a leitura de:

NATIONAL ARCHIVES OF AUSTRALIA. *Developing a functions thesaurus: guidelines for Commonwealth Agencies*. Canberra, 2000.

ARCHIVES AUTHORITY OF NEW SOUTH WALES. *The keyword AAA: a thesaurus of general terms*. Sydney, 1998.

NATIONAL ARCHIVES OF AUSTRALIA. *Designing and implementing recordkeeping systems (DIRKS) manual*. Canberra, 2000.

Todas as publicações editadas pelo Arquivo Nacional da Austrália podem ser acessadas a partir do site <http://www.naa.gov.au/recordkeeping>.

Retiramos exemplos, e inspiração, dos seguintes vocabulários controlados utilizados em sistemas arquivísticos:

ACHIAMÉ, F.A.M. *Plano de classificação funcional de documentos*. Vitória: Estado do Espírito Santo/Secretaria de Estado da Fazenda, 1994.

DOMÈNECH i FERNÁNDEZ, S. *Tesaure BIMA*. Barcelona: Arxiu Municipal de Barcelona, [1993].

ESTADO DO ESPÍRITO SANTO. Tribunal de Contas. *Plano de classificação geral*. Vitória, 1999.

ESTADO DO PARANÁ. Departamento Estadual de Arquivo Público. *Manual de gestão de documentos do Estado do Paraná*. 2ª ed. Curitiba, 1998.

ESTADO DO RIO DE JANEIRO. Secretaria de Estado de Administração e Reestruturação. *Manual de procedimentos para classificação de documentos do poder executivo*. Rio de Janeiro, 2001. V.1: Atividades-meio.

ROBERGE, M. *La classification universelle des documents administratifs*. Québec: Documentor, 1985.

SISTEMA DE ARQUIVOS DA UNIVERSIDADE DE SÃO PAULO. *Plano de classificação das atividades da Universidade de São Paulo*. São Paulo, 1997.

TATEM, J.M., ROLLISON, J. *Thesaurus of university terms developed at Case Western Reserve University Archives*. Chicago: Society of American Archivists, [1985].

UNIVERSIDAD CARLOS III DE MADRID. *Cuadro de clasificación de los documentos*. Madri, 1996.

De resto, adaptamos os procedimentos para controle de vocabulário a partir de uma bibliografia biblioteconômica:

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 6034: Preparação de índice de publicações*. Rio de Janeiro, 1989.

AUSTIN, D. *Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües*. Brasília: IBICT, SENAI, 1993.

CAVALCANTI, C.R. *Indexação & tesouro: metodologia & técnicas*. Ed. preliminar. Brasília: Associação de Bibliotecários do Distrito Federal, 1978.

CINTRA, A.M.M. et al. Linguagens documentárias e terminologia. In: ALVES, I.M. (org.). *A constituição da normalização terminológica no Brasil*. São Paulo: FFLCH/CITRAT, 1996. (Cadernos de terminologia, 1). p.17-22.

CINTRA, A.M.M. et al. *Para entender as linguagens documentárias*. São Paulo: Polis, APB, 1994. (Palavra Chave, 4).

GOMES, H.E. *Diretrizes para elaboração de tesouros monolíngües*. Brasília: IBICT, 1984.

GOMES, H.E. *Manual de elaboração de tesouros monolíngües*. Brasília: Programa Nacional de Bibliotecas das Instituições de Ensino Superior, 1990.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *ISO 2788: Documentation guidelines for the establishment and development of monolingual thesauri*. 1986.

LANCASTER, F.W. *Construção e uso de tesouros: curso condensado*. Brasília: IBICT, 1987.

SMIT, J.W. (org.). *Análise documentária: a análise da síntese*. Brasília: IBICT, 1987.

TÁLAMO, M.F.G.M. et al. A contribuição da terminologia para a elaboração de tesouros. *Ciência da Informação*, Brasília, v.21, n.3, p.197-200, 1993.

## ANEXO

Ficha-síntese das abreviaturas (inglês e português) utilizadas para designar as relações entre os termos nos vocabulários controlados.

Relação entre os termos	Plano de Classificação	Tesouro	Explicitação das siglas utilizadas nos tesouros
Empregar	VER USE SEE	USE	
Empregado para	-	UP UF	Usado para Used for
Termo genérico	-	TG BT	Termo genérico Broader term
Termo específico	-	TE NT	Termo específico Narrower term
Termo relacionado	VER TAMBÉM SEE ALSO	TR RT	Termo relacionado Related term
Nota de explicação ou uso	-	NE SN NU	Nota de escopo Scope note Nota de uso