

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
ESCOLA DE COMUNICAÇÃO  
MESTRADO EM CIÊNCIA DA INFORMAÇÃO

RECUPERAÇÃO DE INFORMAÇÕES EM LIVROS E MONOGRAFIAS  
algoritmo conceitual e prático para maior revocação

por

Francisca Maria de Mello Leal Santiago Lamas

v.1

Dissertação apresentada ao  
CNPq/IBICT e a UFRJ para a  
obtenção do grau de Mestre em  
Ciência da Informação.

Orientadora: Gilda Maria Braga,  
Ph.D., CNPq/IBICT.

Rio de Janeiro

1992

## AGRADECIMENTOS

Ao BANCO NACIONAL DE DESENVOLVIMENTO ECONÔMICO E SOCIAL - BNDES, pela oportunidade de cursar o Mestrado em Ciência da Informação IBICT/UFRJ.

Ao CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq, pelo apoio financeiro concedido para a realização desta Pesquisa.

À Profª Gilda Maria Braga, pela constante e valiosa orientação.

Ao Corpo Docente do Mestrado em Ciência da Informação, em especial à Profª Heloísa Tardin Christovão, pelos conhecimentos apreendidos e pela dedicação para com todos os seus alunos.

Ao pessoal da Biblioteca da Escola de Comunicação da UFRJ, da Biblioteca do BNDES, da Biblioteca da Escola Superior de Guerra - ESG e da Rede BIBLIODATA (CALCO), pelas informações e material bibliográfico obtidos, essenciais à elaboração da presente Pesquisa.

A todos aqueles, principalmente meus familiares, que me apoiaram e acreditaram na realização deste trabalho.

Muito obrigada

# SUMÁRIO

## Primeiro Volume

1. INTRODUÇÃO p.1
2. O PROCESSAMENTO TÉCNICO DE LIVROS E MONOGRAFIAS NOS SISTEMAS DE INFORMAÇÃO p.14
  - 2.1 - Catologação e Política de Indexação p.16
  - 2.2 - Alterações necessárias no processamento técnico de livros e monografias p.20
3. REVISÃO DA LITERATURA p.25
4. OBJETIVOS p.40
5. JUSTIFICATIVA p.42
6. MATERIAL p.47
  - 6.1 - A identificação da amostra p.47
  - 6.2 - O tamanho da amostra p.48
7. METODOLOGIA p.54
  - 7.1 - Definição dos elementos estruturais de livros e monografias a utilizar na análise conceitual p.54
    - 7.1.1 - Os títulos, subtítulos e sumários p.58
  - 7.2 - Análise conceitual p.62
    - 7.2.1 - Seleção de palavras e/ou expressões p.64
    - 7.2.2 - Registro da análise conceitual p.70
  - 7.3 - Compatibilização dos cabeçalhos de assunto formados a partir da análise conceitual com os vocabulários controlados adotados pelos Sistemas de Informação investigados p.72
    - 7.3.1 - Na Biblioteca do BNDES p.72
      - 7.3.1.1 - Caracterização dos cabeçalhos formados p.73
      - 7.3.2 - Na Rede BIBLIODATA (CALCO) p.76
        - 7.3.2.1 - Caracterização dos cabeçalhos formados p.77
    - 7.4 - Estudo da distribuição, por páginas de texto, dos níveis de subordinação dos sumários dos documentos integrantes da amostra p.81

8. RESULTADOS p.90

- 8.1 - Resultados relativos à pós-coordenação p.94
- 8.1.1 - Acréscimos obtidos junto à Biblioteca do BNDES p.96
- 8.1.2 - Potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões destinadas à indexação de livros e monografias em Sistemas de Informação pós-coordenados p.115
- 8.1.2.1 - Os cabeçalhos cronológicos p.116
- 8.1.2.2 - Os cabeçalhos não disponíveis no vocabulário controlado p.118
- 8.1.2.3 - Os cabeçalhos coincidentes p.126
- 8.1.2.4 - O potencial dos títulos, subtítulos e sumários p.129
- 8.2 - Resultados relativos à pré-coordenação p.139
- 8.2.1 - Acréscimos obtidos junto à Rede BIBLIODATA (CALCO) p.141
- 8.2.2 - Acréscimos obtidos junto à Base de Dados bibliográficos da Rede BIBLIODATA (CALCO) p.161
- 8.2.2.1 - Acréscimos sem os subcabeçalhos cronológicos p.162
- 8.2.2.2 - Acréscimos com os subcabeçalhos cronológicos p.166
- 8.2.3 - Potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões destinadas à indexação de livros e monografias em Sistemas de Informação pré-coordenados p.175
- 8.2.3.1 - O potencial na Rede BIBLIODATA (CALCO) p.176
- 8.2.3.1.1 - Os cabeçalhos simples e compostos não disponíveis no vocabulário controlado p.177
- 8.2.3.1.2 - Os cabeçalhos simples e compostos coincidentes 183
- 8.2.3.1.3 - A demonstração do potencial na Rede BIBLIODATA (CALCO) p.185
- 8.2.3.2 - O potencial na Base de Dados bibliográficos da Rede BIBLIODATA (CALCO) p.194
- 8.2.3.2.1 - A demonstração do potencial sem os subcabeçalhos cronológicos p.198
- 8.2.3.2.2 - A demonstração do potencial com os subcabeçalhos cronológicos p.205
- 8.3 - Comparação dos resultados obtidos para a pós-coordenação e para a pré-coordenação p.210
- 8.4 - Relações entre a quantificação dos cabeçalhos de assunto formados e a quantificação das páginas de texto correspondentes aos níveis de subordinação dos sumários dos documentos integrantes da amostra examinada p.214

9. CONCLUSÕES p.224

10. REFERÊNCIAS BIBLIOGRÁFICAS p.234



*Segundo Volume*

**ANEXOS**

- Anexo 01 - DESCRIÇÃO BIBLIOGRÁFICA, CABEÇALHOS DE ASSUNTO ATRIBUÍDOS PELOS SISTEMAS DE INFORMAÇÃO INVESTIGADOS E LOCALIZAÇÃO NOS RESPECTIVOS ACERVOS DOS DOCUMENTOS INTEGRANTES DA AMOSTRA p.1
- Anexo 02 - CÓPIAS DOS SUMÁRIOS DOS DOCUMENTOS INTEGRANTES DA AMOSTRA p.19
- |              |      |
|--------------|------|
| Documento 01 | p.19 |
| Documento 02 | p.23 |
| Documento 03 | p.27 |
| Documento 04 | p.30 |
| Documento 05 | p.31 |
| Documento 06 | p.32 |
| Documento 07 | p.34 |
| Documento 08 | p.36 |
| Documento 09 | p.37 |
| Documento 10 | p.38 |
| Documento 11 | p.39 |
| Documento 12 | p.40 |
| Documento 13 | p.41 |
| Documento 14 | p.42 |
| Documento 15 | p.43 |
| Documento 16 | p.44 |
| Documento 17 | p.45 |
| Documento 18 | p.46 |
| Documento 19 | p.47 |
| Documento 20 | p.49 |
| Documento 21 | p.51 |
| Documento 22 | p.52 |
| Documento 23 | p.53 |
| Documento 24 | p.54 |
| Documento 25 | p.55 |
| Documento 26 | p.56 |
| Documento 27 | p.57 |
| Documento 28 | p.58 |
| Documento 29 | p.59 |
| Documento 30 | p.60 |
- Anexo 03 - NÍVEIS MÁXIMOS DE SUBORDINAÇÃO CONFERIDOS AOS SUMÁRIOS DOS DOCUMENTOS INTEGRANTES DA AMOSTRA p.61
- Anexo 04 - ANÁLISE CONCEITUAL: QUANTIFICAÇÃO DAS PALAVRAS, EXPRESSÕES E NOMES PRÓPRIOS SELECIONADOS NOS DOCUMENTOS INTEGRANTES DA AMOSTRA, POR NÍVEL DE SUBORDINAÇÃO p.62
- Anexo 05 - ANÁLISE CONCEITUAL: DISCRIMINAÇÃO DAS PALAVRAS, EXPRESSÕES E NOMES PRÓPRIOS SELECIONADOS NOS TÍTULOS, SUBTÍTULOS E SUMÁRIOS, PARA CADA UM DOS DOCUMENTOS INTEGRANTES DA AMOSTRA, NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO p.63

- Anexo 06 - BIBLIOTECA DO BNDES: DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO FORMADOS A PARTIR DA ANÁLISE CONCEITUAL, PARA CADA UM DOS DOCUMENTOS INTEGRANTES DA AMOSTRA, POR NÍVEL DE SUBORDINAÇÃO p. 79
- Anexo 07 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO FORMADOS A PARTIR DA ANÁLISE CONCEITUAL, PARA CADA UM DOS DOCUMENTOS INTEGRANTES DA AMOSTRA, POR NÍVEL DE SUBORDINAÇÃO p. 94
- Anexo 08 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO ACRESCIDOS PELA INVESTIGAÇÃO E ASSOCIAÇÕES CORRESPONDENTES p. 108
- Anexo 09 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS CRONOLÓGICOS p. 117
- Anexo 10 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO p. 121
- Anexo 11 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO COINCIDENTES p. 125
- Anexo 12 - BIBLIOTECA DO BNDES: DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO E ASSOCIAÇÕES CORRESPONDENTES ACRESCIDOS, POR DOCUMENTO p. 134
- Anexo 13 - QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO E ASSOCIAÇÕES CORRESPONDENTES ATRIBUÍDOS PELA BIBLIOTECA DO BNDES AOS DOCUMENTOS INTEGRANTES DA AMOSTRA p. 147
- Anexo 14 - BIBLIOTECA DO BNDES: COMPARAÇÃO ENTRE AS MÉDIAS DE CABEÇALHOS E ASSOCIAÇÕES ATRIBUÍDOS PELA BIBLIOTECA E AS MÉDIAS DE CABEÇALHOS E ASSOCIAÇÕES ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO p. 149
- Anexo 15 - BIBLIOTECA DO BNDES: DISCRIMINAÇÃO DOS CABEÇALHOS CRONOLÓGICOS, POR DOCUMENTO p. 150
- Anexo 16 - BIBLIOTECA DO BNDES: DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR DOCUMENTO p. 153
- Anexo 17 - BIBLIOTECA DO BNDES: ORDENAÇÃO ALFABÉTICA DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO p. 158
- Anexo 18 - BIBLIOTECA DO BNDES: DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO E ASSOCIAÇÕES CORRESPONDENTES COINCIDENTES, POR DOCUMENTO p. 163
- Anexo 19 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO ACRESCIDOS PELA INVESTIGAÇÃO p. 166

- Anexo 20 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO p.176
- Anexo 21 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO COINCIDENTES p.186
- Anexo 22 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS ACRESCIDOS, POR DOCUMENTO p.196
- Anexo 23 - QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO ATRIBUÍDOS PELA REDE BIBLIODATA (CALCO) AOS DOCUMENTOS INTEGRANTES DA AMOSTRA, COM DETALHAMENTO DE PONTOS DE ACESSO À BASE DE DADOS DA REDE p.205
- Anexo 24 - REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE AS MÉDIAS DE CABEÇALHOS DE ASSUNTO ATRIBUÍDOS PELA REDE E AS MÉDIAS DE CABEÇALHOS DE ASSUNTO ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO p.207
- Anexo 25 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS QUE COMPUSERAM A FORMAÇÃO DOS CABEÇALHOS COMPOSTOS ACRESCIDOS, POR DOCUMENTO p.208
- Anexo 26 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS SUBCABEÇALHOS CRONOLÓGICOS ACRESCIDOS, POR DOCUMENTO p.214
- Anexo 27 - REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE AS MÉDIAS DE PONTOS DE ACESSO DISPONÍVEIS NA BASE DE DADOS BIBLIOGRÁFICOS DA REDE E AS MÉDIAS DE PONTOS DE ACESSO ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO p.217
- Anexo 28 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR DOCUMENTO p.218
- Anexo 29 - REDE BIBLIODATA (CALCO): ORDENAÇÃO ALFABÉTICA DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO p.224
- Anexo 30 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS DE ASSUNTO COINCIDENTES, POR DOCUMENTO p.230
- Anexo 31 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES DOS CABEÇALHOS COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR DOCUMENTO p.232
- Anexo 32 - REDE BIBLIODATA (CALCO): DISCRIMINAÇÃO DOS CABEÇALHOS TÓPICOS, SUBCABEÇALHOS E SUBCABEÇALHOS CRONOLÓGICOS INTEGRANTES DOS CABEÇALHOS COMPOSTOS COINCIDENTES, POR DOCUMENTO p.236

RESUMO

Anexo 33 - CONTAGEM DAS PÁGINAS DE TEXTO CORRESPONDENTES AOS NÍVEIS DE SUBORDINAÇÃO DEFINIDOS PARA OS SUMÁRIOS DOS DOCUMENTOS INTEGRANTES DA AMOSTRA p.237

Este documento apresenta os resultados da contagem das páginas de texto correspondentes aos níveis de subordinação definidos para os sumários dos documentos integrantes da amostra. O estudo foi realizado com o objetivo de verificar a aderência dos documentos à estrutura de subordinação estabelecida no manual de redação. Os resultados são apresentados em tabelas e gráficos, permitindo a análise da distribuição das páginas em cada nível de subordinação. O estudo também identificou algumas inconsistências e sugeriu melhorias para a estrutura dos sumários. Os resultados são apresentados em tabelas e gráficos, permitindo a análise da distribuição das páginas em cada nível de subordinação. O estudo também identificou algumas inconsistências e sugeriu melhorias para a estrutura dos sumários.

## RESUMO

Aumento da acessibilidade ao conteúdo informativo de livros e monografias, através de indexações mais exaustivas, sem comprometer os níveis de precisão a ocorrerem na recuperação de informações. A metodologia empregada previu a utilização dos títulos, subtítulos e sumários de livros e monografias a fim de se identificar palavras com elevada capacidade de representação conceitual, destinadas à posterior formação de cabeçalhos de assunto ou termos de indexação. São examinados livros e monografias que igualmente integram os acervos de dois Sistemas de Informação brasileiros: um adepto da pós-coordenação, para os processos de indexação e de recuperação de informações, e outro seguidor da pré-coordenação, respectivamente, a Biblioteca do Banco Nacional de Desenvolvimento Econômico e Social - BNDES e a Rede BIBLIODATA (CALCO). Dos títulos, subtítulos e sumários destes documentos foram selecionadas palavras, expressões e nomes próprios para a formação de cabeçalhos pertinentes aos dois Sistemas. Nos sumários dos documentos foram definidos níveis de subordinação, de acordo com sua estrutura hierárquica. Os resultados da pesquisa revelaram substanciais acréscimos em relação à catalogação por assunto efetuada pelos dois Sistemas, como também evidenciaram o potencial dos títulos, subtítulos e sumários para a identificação de palavras com elevado grau de representatividade conceitual dos documentos. Os resultados também demonstraram que, recorrendo-se aos títulos, subtítulos e até ao nível 2 de subordinação dos sumários de livros e monografias com menor especificidade de tratamento dos assuntos, e até ao nível 3 de subordinação daqueles com maior especificidade, tanto em Sistemas de Informação pós-coordenados quanto em Sistemas pré-coordenados, a formação de cabeçalhos ou de termos de indexação atinge níveis exaustivos e consistentes, com representatividade conceitual pertinente para a posterior recuperação das informações veiculadas pelos documentos. A pesquisa também constatou deficiências nas linguagens de indexação utilizadas pelos dois Sistemas, como também as limitações de acesso ao conteúdo informativo dos documentos examinados decorrentes do reduzido número de cabeçalhos ou termos de indexação que lhes foram atribuídos pelos Sistemas. Sugere-se a revisão destas linguagens e das Políticas de Indexação vigentes nos Sistemas, objetivando o aprimoramento da qualidade de seus serviços.

## ABSTRACT

Improvement of subject access to books and monographies, through more exhaustive indexing processes, without substantially decreasing information retrieval precision. The applied methodology used titles, subtitles and tables of contents of books and monographies to identify words with a strong capability of conceptual representation for additional subject headings or indexing terms formulation. The research consisted of the analysis of books and monographies included in the bibliographical collection of two different Brazilian Information Systems: a coordinated system and a pre-coordinated one, respectively, the library of the Banco Nacional de Desenvolvimento Econômico e Social - BNDES, and the BIBLIODATA (CALCO) Network. From the titles, subtitles and tables of contents of these documents, words, expressions and proper names were selected for the construction of appropriate subject headings. Hierarchical levels were established to documents' tables of contents, according to their structures. Research results have led to substantial increases in subject access when compared to subject cataloguing made by the two above-mentioned Information Systems. Results demonstrated the potential of titles, subtitles and tables of contents for identifying words with a high degree of documents' conceptual representation. The results have also demonstrated that the use of words extracted from titles, subtitles and until the second hierarchical level of tables of contents of books and monographies with a less specific approach of subjects, and down to the third hierarchical level of those documents with a more specific approach, the subject headings or indexing terms formulation achieves exhaustive and consistent levels, with a great and pertinent conceptual representativeness for future information retrieval. Additionally, the research has pointed out deficiencies in the indexing languages used by both examined Information Systems, and has confirmed the limitations of subject access to books and monographies due to the reduced number of subject headings assigned by these Systems. A careful analysis of both indexing languages and indexing policies used by the two Information Systems is suggested to improve the quality of their services.

## LISTA DE QUADROS

- Quadro 01 - REDE BIBLIODATA (CALCO): CODIFICAÇÕES ATRIBUÍDAS AOS CABEÇALHOS DE ASSUNTO FORMADOS *p. 80*
- Quadro 02 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO ACRESCIDOS E ASSOCIAÇÕES CORRESPONDENTES IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS *p. 97*
- Quadro 03 - BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS ACRESCIDOS E ASSOCIAÇÕES CORRESPONDENTES, POR GRUPOS DE DOCUMENTOS *p. 98*
- Quadro 04 - BIBLIOTECA DO BNDES: CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS ACRESCIDOS E RESPECTIVAS ASSOCIAÇÕES, POR GRUPOS DE DOCUMENTOS *p. 99*
- Quadro 05 - BIBLIOTECA DO BNDES: PERCENTUAIS DE CABEÇALHOS ACRESCIDOS E ASSOCIADOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS *p. 102*
- Quadro 06 - BIBLIOTECA DO BNDES: ACRÉSCIMOS À CATALOGAÇÃO POR ASSUNTO OBTIDOS COM OS CABEÇALHOS E ASSOCIAÇÕES CORRESPONDENTES FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS *p. 104*
- Quadro 07 - BIBLIOTECA DO BNDES: VALORES ACUMULADOS, NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, PARA OS CABEÇALHOS ACRESCIDOS E ASSOCIAÇÕES CORRESPONDENTES, POR GRUPOS DE DOCUMENTOS *p. 106*
- Quadro 08 - BIBLIOTECA DO BNDES: PERCENTUAIS DE ACRÉSCIMOS CUMULATIVOS OBTIDOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS *p. 107*
- Quadro 09 - BIBLIOTECA DO BNDES: DIFERENÇAS DE PERCENTUAIS DE ACRÉSCIMOS ENTRE OS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS *p. 109*
- Quadro 10 - BIBLIOTECA DO BNDES: COMPARAÇÃO ENTRE AS MÉDIAS DE CABEÇALHOS E ASSOCIAÇÕES ATRIBUÍDOS PELA BIBLIOTECA E AS MÉDIAS DE CABEÇALHOS E ASSOCIAÇÕES ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO *p. 113*
- Quadro 11 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS CRONOLÓGICOS IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS *p. 116*
- Quadro 12 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS *p. 119*

- Quadro 13 - BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS p.120
- Quadro 14 - BIBLIOTECA DO BNDES: CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS p.121
- Quadro 15 - BIBLIOTECA DO BNDES: PERCENTUAIS DE CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOC. CONTROLADO, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS p.123
- Quadro 16 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO COINCIDENTES E ASSOCIAÇÕES CORRESPONDENTES IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS p.126
- Quadro 17 - BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS COINCIDENTES E ASSOCIAÇÕES CORRESPONDENTES, POR GRUPOS DE DOCUMENTOS p.128
- Quadro 18 - BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS p.130
- Quadro 19 - BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, POR DOCUMENTO p.131
- Quadro 20 - BIBLIOTECA DO BNDES: CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS p.132
- Quadro 21 - BIBLIOTECA DO BNDES: PERCENTUAIS DE CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, ACUMULADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS p.133
- Quadro 22 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS p.142
- Quadro 23 - REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS ACRESCIDOS, POR DOCUMENTO p.143
- Quadro 24 - REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS SIMPLES E COMPOSTOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS p.145
- Quadro 25 - REDE BIBLIODATA (CALCO): PERCENTUAIS DE CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS ACRESCIDOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS p.148



- Quadro 26 - REDE BIBLIODATA (CALCO): ACRÉSCIMOS À CATALOGAÇÃO POR ASSUNTO OBTIDOS COM OS CABEÇALHOS SIMPLES E COMPOSTOS FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS *p. 151*
- Quadro 27 - REDE BIBLIODATA (CALCO): VALORES ACUMULADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO PARA OS CABEÇALHOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS *p. 153*
- Quadro 28 - REDE BIBLIODATA (CALCO): PERCENTUAIS DE ACRÉSCIMOS CUMULATIVOS OBTIDOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS *p. 154*
- Quadro 29 - REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE AS MÉDIAS DE CABEÇALHOS ATRIBUÍDOS PELA REDE E AS MÉDIAS DE CABEÇALHOS ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO *p. 160*
- Quadro 30 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES DOS CABEÇALHOS COMPOSTOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS *p. 162*
- Quadro 31 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS SIMPLES E COMPOSTOS ACRESCIDOS E DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES, POR GRUPOS DE DOCUMENTOS *p. 163*
- Quadro 32 - REDE BIBLIODATA (CALCO): ACRÉSCIMOS À BASE DE DADOS BIBLIOGRÁFICOS, SEM OS SUBCABEÇALHOS CRONOLÓGICOS, POR GRUPOS DE DOCUMENTOS *p. 164*
- Quadro 33 - REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE OS ACRÉSCIMOS OBTIDOS PARA A REDE E PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS *p. 165*
- Quadro 34 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS SUBCABEÇALHOS CRONOLÓGICOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS *p. 167*
- Quadro 35 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DE TODOS OS CABEÇALHOS E SUBCABEÇALHOS ACRESCIDOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS *p. 168*
- Quadro 36 - REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS E SUBCABEÇALHOS ACRESCIDOS À BASE DE DADOS, POR DOCUMENTO *p. 169*
- Quadro 37 - REDE BIBLIODATA (CALCO): ACRÉSCIMOS TOTAIS À BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS *p. 170*

- Quadro 38 - REDE BIBLIODATA (CALCO): VALORES ACUMULADOS, NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, PARA OS CABEÇALHOS E SUBCABEÇALHOS ACRESCIDOS À BASE DE DADOS, POR GRUPOS DE DOCUMENTOS p.171
- Quadro 39 - REDE BIBLIODATA (CALCO): PERCENTUAIS DE ACRÉSCIMOS CUMULATIVOS OBTIDOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO PARA A BASE DE DADOS, POR GRUPOS DE DOCUMENTOS p.172
- Quadro 40 - REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE AS MÉDIAS DE PONTOS DE ACESSO À BASE DE DADOS ATRIBUÍDOS PELA REDE E AS MÉDIAS DE PONTOS DE ACESSO ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO p.174
- Quadro 41 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS SIMPLES E COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS p.177
- Quadro 42 - REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS SIMPLES E COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR DOCUMENTO p.179
- Quadro 43 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO E PERCENTUAIS DE CABEÇALHOS SIMPLES E COMPOSTOS NÃO DISPONÍVEIS NO VOC. CONTROLADO, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS p.180
- Quadro 44 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS SIMPLES E COMPOSTOS COINCIDENTES IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS p.184
- Quadro 45 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS SIMPLES E COMPOSTOS ACRESCIDOS, NÃO DISPONÍVEIS NO VOC. CONTROLADO E COINCIDENTES, POR GRUPOS DE DOCUMENTOS p.186
- Quadro 46 - REDE BIBLIODATA (CALCO): MÉDIAS DO TOTAL DE CABEÇALHOS SIMPLES E COMPOSTOS FORMADOS, POR DOCUMENTO p.187
- Quadro 47 - REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS FORMADOS, POR GRUPOS DE DOCUMENTOS p.188
- Quadro 48 - REDE BIBLIODATA (CALCO): PERCENTUAIS DE CABEÇALHOS DE ASSUNTO FORMADOS, ACUMULADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS p.190
- Quadro 49 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES DOS CABEÇALHOS COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS p.195

- Quadro 50 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES DOS CABEÇALHOS COMPOSTOS COINCIDENTES, POR GRUPOS DE DOCUMENTOS p.197
- Quadro 51 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DO POTENCIAL DOS TÍTULOS, SUBTÍTULOS E SUMÁRIOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, SEM OS SUBCABEÇALHOS CRONOLÓGICOS, POR GRUPOS DE DOCUMENTOS p.199
- Quadro 52 - REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR DOCUMENTO p.200
- Quadro 53 - REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS p.203
- Quadro 54 - REDE BIBLIODATA (CALCO): PERCENTUAIS DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS p.204
- Quadro 55 - REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DO POTENCIAL DOS TÍTULOS, SUBTÍTULOS E SUMÁRIOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, COM OS SUBCABEÇALHOS CRONOLÓGICOS, POR GRUPOS DE DOCUMENTOS p.206
- Quadro 56 - REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS (INCLUINDO OS CRONOLÓGICOS), PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS p.207
- Quadro 57 - REDE BIBLIODATA (CALCO): PERCENTUAIS DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO (INCLUINDO OS SUBCABEÇALHOS CRONOLÓGICOS), PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS p.208

## 1. INTRODUÇÃO

A evolução da sociedade tem se manifestado sob diversas formas e, de algum modo, mesmo que primitivo, as transformações sociais sempre foram registradas e disseminadas pelo homem.

No que se refere à evolução do aprendizado, do conhecimento, do saber, desde o período pré-histórico o homem procurou deixar registrado o que talvez supunha ser de interesse e/ou utilização coletiva para as gerações do momento e posteriores, atitude que, comprovadamente, é inerente à natureza humana.

O meio formal mais antigo de apresentação e de disseminação desse registro entre povos e civilizações é o livro.

É difícil precisar cronologicamente quando o homem começou a gerar e a produzir livros, devido ao estado de conservação dos materiais na época empregados para a confecção dos exemplares mais antigos descobertos e até então preservados (01).

Já a partir da Idade Antiga, os Sumerianos, Babilônios e Assírios escreviam livros em placas de argila. Os Egípcios utilizavam os rolos de papiro, um material muito semelhante ao papel, que, embora já descoberto pelos chineses, ainda não era conhecido pelas culturas da bacia do Mediterrâneo e da Ásia Menor. Os Gregos e Romanos também adotaram o rolo de papiro para a confecção de seus livros. Por sua vez, os Hebreus, Gregos e

Romanos desenvolveram o uso do pergaminho, um novo material destinado ao assentamento da escrita, ainda na forma de rolos, derivado do couro de cabras, ovelhas ou vitelas, material esse mais resistente, flexível e duradouro. Mais tarde, o codex veio revolucionar a apresentação dos livros, pois ao invés do papiro ou do pergaminho ser enrolado, folhas soltas dos mesmos materiais eram reunidas na seqüência textual de forma similar às atuais encadernações. Por volta do ano 400, o codex de pergaminho tornou-se a forma dominante de transmissão da palavra escrita no Ocidente (02).

A dissolução do Império Romano (Século V) e o conseqüente domínio dos bárbaros comprometeram a existência dos livros. Felizmente, na ocasião, a Igreja resistiu aos freqüentes assaltos, tornando-se a instituição detentora e preservadora do conhecimento registrado pelos livros. Eram os mosteiros medievais que os confeccionavam, adornavam e preservavam. As Bibliotecas monásticas desempenharam um importante papel na transmissão do conhecimento, apesar das restrições de acesso que impunham ao público externo (03).

A descoberta da imprensa por Gutenberg no Século XV, assim como a chegada do papel no mundo ocidental, alteraram radicalmente a forma dos livros e os tornaram mais acessíveis para os povos interessados.

A mecanização da imprensa, a extensão do estudo e da cultura a outras classes sociais, e a modificação dos regimes políticos, entre outros fatores, ocasionaram uma maior disponibilidade de informações registradas, durante a Idade

Moderna. Por essa razão, aumentou o número de Bibliotecas públicas e especializadas, estas geralmente vinculadas a Instituições de Pesquisa. O livro tornou-se sinônimo de instrumento transmissor da informação sedimentada nos meios culturais e acadêmicos.

A intensidade da evolução científica e tecnológica da sociedade ocorrida a partir do Século XVIII alterou as características e os propósitos de alguns livros, tornando-os mais especializados e dirigidos a um público específico, em função das frequentes descobertas e aplicações surgidas e divulgadas por cientistas e pesquisadores.

Acompanhando as principais evoluções registradas na história dos povos e civilizações, o livro se consolidou como o documento mais popular e acessível para a transmissão e o aprendizado do conhecimento científico, cultural, social e tecnológico.

Entretanto, as grandes descobertas científicas e tecnológicas ocorridas a partir da Revolução Industrial (Século XVII) provocaram a necessidade de se disseminar as informações especializadas mais rapidamente. Surgiram, então, os periódicos científicos, geralmente editados por Academias e instituições afins, cuja finalidade primordial era informar restritamente a seus membros a evolução e o acompanhamento das atividades de pesquisa efetuadas por seus pares.

A partir do início do Século XX, a edição de periódicos especializados, principalmente aqueles de cunhos científico e tecnológico, atingiu números em proporções já surpreendentes (04), seguramente ocasionada pela crescente necessidade de se disseminar informações com caráter mais imediatista, em função da competitividade e das manifestações de soberania provocadas pelo desenvolvimento científico e tecnológico das nações então estabelecidas. Também o aumento dos níveis de especialização do conhecimento contribuiu para o lançamento de um número cada vez maior de títulos de periódicos especializados.

De certa forma, o periódico tornou-se o principal componente informativo dos acervos das Bibliotecas e Centros de Informação, sobretudo daqueles especializados, superando os livros quanto à atualização e à velocidade de transmissão de informações textuais.

É sabida a dimensão do universo compreendido pelos periódicos especializados e a importância que lhes é conferida pelos Sistemas de Informação atuais.

Essa inversão de valores ocorrida entre o livro e o periódico especializado no processo de transmissão de informações seguramente orientou as primeiras pesquisas no âmbito da Ciência da Informação. Desde sua origem, fato que ocorreu por volta de 1945 (05), observa-se na literatura produzida pela Ciência da Informação que existe uma maior concentração de investigações sobre o periódico como suporte físico de veiculação da informação especializada, como também o

inverso, ou seja, pesquisas sobre determinada literatura assentada no periódico especializado. Diversos estudos têm sido formulados, principalmente de natureza empírica. Essa manifestação pode ser atribuída às principais características da informação contida nos periódicos (curto espaço de tempo entre sua formulação e sua edição, e a atualização de seu conteúdo) e à importância que grande parte dos usuários dos Sistemas de Informação confere a esses documentos. O periódico especializado divulga formalmente a seu público alvo, em curto prazo, os mais recentes resultados de pesquisas, tendências e diretrizes de determinadas investigações, implementação de novos métodos e processos, etc. Entretanto, devido à intensidade e à variedade de edição de artigos e às frequentes inovações e aperfeiçoamentos ocorridas nos métodos, processos e técnicas, raros são os artigos de periódicos especializados que não atingem um relativo índice de obsolescência em curto espaço de tempo. Esses índices de obsolescência podem variar não somente de acordo com os propósitos informacionais e com o padrão de qualidade de um periódico para determinada literatura, mas também de acordo com a intensidade de uso de seus artigos no processo de comunicação formal (06).

Um outro dado atinente aos periódicos especializados refere-se ao investimentos efetuados pelos Sistemas de Informação em seu processamento técnico, principalmente aqueles relativos ao desenvolvimento e aperfeiçoamento de técnicas de indexação, de disseminação e de recuperação de informações. Verifica-se, atualmente, a existência de um grande número de serviços especializados (Bases de Dados Internacionais e Nacionais e Serviços de Resumo, como exemplo), cuja maior parte



dos documentos que processa é constituída por periódicos especializados. Esses serviços reportam-se principalmente à Ciência e à Tecnologia.

Por outro lado, grandes transformações ocorreram a partir da segunda metade do Século XX quanto às características físicas e informacionais dos documentos denominados livros. Independente da disponibilidade do periódico especializado, a intensificação da especialização do conhecimento científico, tecnológico e social também provocou a edição de estudos extensos e minuciosos, geralmente individuais ou produzidos por Instituições públicas e privadas (como exemplos, Dissertações, Teses e Relatórios Técnicos ou Científicos), que se propunham esgotar um determinado tema específico, sem, entretanto, possuírem cunho comercial. Esses documentos, as monografias, apresentados através de editoração convencional, mimeografia, xerografia ou, como nos dias de hoje, de "softwares" de editores de texto disponíveis em microprocessadores, vieram acrescentar à comunicação formal um novo e importante meio de transmissão da informação especializada.

É oportuno mencionar que livros e monografias, conceitualmente, possuem características e propósitos distintos, mas, por uma questão de terminologia, em alguns momentos, são considerados como categorias idênticas de documentos.

Retornando ao âmbito da Ciência da Informação, e das ciências interdisciplinares, como a Biblioteconomia, poucos são os estudos e pesquisas dedicados à informação veiculada por livros e monografias, ainda mais no que se refere à análise e ao

maior aproveitamento de seu conteúdo informativo.

Os livros e as monografias, mesmo possuindo características e objetivos informacionais diferentes dos periódicos, principalmente aqueles referentes à sedimentação de um determinado campo do conhecimento, ao tempo dispendido durante a sua elaboração, e à divulgação de informações que venham a prescrever a médio e longo prazos, nem por isso devem ser preteridas pelo periódico quanto a estudos e pesquisas de sua participação no processo de transmissão do conhecimento. Ainda nos dias de hoje, livros e monografias, em especial os livros, são documentos que possuem uma expressiva participação percentual nos acervos das Bibliotecas e dos Centros de Documentação públicos e especializados, sobretudo nas áreas de Ciências Humanas e Sociais, sem, entretanto, merecerem por parte dos estudiosos o interesse e um número de pesquisas condizente com a importância e com o nível de seu potencial informativo.

Infelizmente, a participação secular dos livros na disseminação do conhecimento, que conferiu-lhes uma significativa popularidade dentre os meios de transmissão da informação textual (superados apenas pelos jornais diários), foi postergada pelo predomínio da utilização do periódico especializado, principalmente os periódicos de cunho científico e tecnológico, fato que, por certo, direcionou grande parte das pesquisas em Ciência da Informação.

Entretanto, faz-se oportuno citar um estudo bibliométrico elaborado por pesquisadores do Institute of Economic Research da Hitotsubashi University, Japão, editado em 1989, o qual enfatiza

a importância dos livros e monografias para a disseminação de informações no âmbito das Ciências Humanas e Sociais, particularmente para a Economia, ressaltando que "... fontes não seriadas constituem ainda as principais citações ..." efetuadas pela literatura sobre esse campo, o que "... faz supor que os livros ainda sejam considerados como muito importantes para os estudiosos dedicados a investigações em Economia ...". Esse mesmo estudo bibliométrico também ressalta a prescrição, em longo prazo, das informações veiculadas por livros e monografias em Ciências Humanas e Sociais (07).

Agora, reportando-se ao contexto mundial, com a instituição de uma nova ordem internacional, a década de 90 veio confirmar as previsões acerca da importância da informação e do domínio do conhecimento como insumos básicos para o desenvolvimento econômico e social. O acesso, a assimilação e o domínio de informações específicas, e até mesmo de caráter estratégico, tornou-se um componente fundamental para o controle de determinadas tecnologias e para a demonstração de poder político e econômico (08). A própria informática, que colocou no mercado equipamentos e "softwares" de uso pessoal e, paralelamente, aprimorou a capacidade de processamento dos computadores de grande porte, veio conferir à sociedade recursos que proporcionaram uma maior velocidade na organização de informações tidas como indispensáveis para a divulgação do conhecimento, para o aprendizado, acompanhamento e detenção das novas tecnologias, e também para o desenvolvimento e consolidação das grandes transformações econômicas e sociais que vêm ocorrendo neste final do Século XX.

Especificamente no caso dos países integrantes do denominado "Terceiro Mundo" (onde se insere o Brasil), para se absorver, assimilar e aplicar as novas tecnologias já à disposição nos países industrializados, ou mesmo criá-las, seria fundamental, dentre outros itens, o preparo educacional de seus povos através de políticas governamentais sérias, realistas e de longo prazo relativas à educação básica e ao treinamento de mão-de-obra qualificada e também especializada (09).

Um dos componentes essenciais do processo educacional são as Bibliotecas e Centros de Documentação, sejam públicos, universitários ou especializados, e, por esta razão, os serviços por eles prestados devem ser cada vez mais aprimorados no sentido de tornar o potencial informativo de seus acervos bibliográficos (geralmente constituídos sob grandes restrições, sobretudo financeiras) mais disponível para o acompanhamento da educação básica ou especializada, como também para o aperfeiçoamento cultural e/ou profissional de seus usuários.

Um outro aspecto a considerar refere-se à acentuada divergência existente entre os padrões de comunicação da informação científica e tecnológica, da informação de caráter social e da informação cultural. No primeiro caso, a comunicação informal supera a formal, prevalecendo a troca de informações pessoais entre pares (10), e quando ocorre o emprego dos meios formais de comunicação, "... os artigos de periódicos são os veículos mais utilizados pela ciência contemporânea ..." (11). Na informação gerada pelas Ciências Humanas e Sociais, existe a expressiva participação dos meios

formais de comunicação, em especial de livros e de monografias (12). A informação de caráter cultural ou de lazer, que interessa a uma minoria das populações, também se sobressai nos meios formais. Assim sendo, pode-se inferir que, com exceção da informação científica e tecnológica, livros e monografias ainda possuem uma expressiva participação no processo de comunicação textual.

No caso da informação de caráter elementar destinada à educação básica, como também para os ensinos colegial e universitário, o principal suporte físico destinado à sua transmissão são os livros e as monografias, e é justamente essa informação que precisa dispor de mecanismos de disseminação e de recuperação mais aprimorados e eficazes para o alcance do desejável preparo dos povos dos países menos favorecidos.

Paralelamente, as inovações tecnológicas relativas ao assentamento e ao armazenamento de informações em meios não convencionais não impediram que o livro e as monografias continuassem a atuar como meios de transmissão de informações, como também os periódicos e os "proceedings". A informação disposta no papel, seja de caráter elementar ou especializado, ainda não se tornou dispensável para o homem, mesmo com a existência de alguns documentos já fixados em suportes eletromagnéticos e com a "informatização" de rotinas administrativas e técnicas.

Todas as considerações anteriormente apresentadas reforçam a necessidade de se continuar investindo maciçamente na informação textual - através de pesquisas teóricas ou empíricas,

ou também através da implementação de processos eficazes destinados à sua acessibilidade - como insumo indispensável para o alcance do equilíbrio político, social e econômico das nações vigentes.

CITAÇÕES E NOTAS

- 01 - ENCYCLOPAEDIA Britannica. Chicago: Encyclopaedia Britannica Inc., 1964. 23 v. v. 3: Balfour to Both.
- 02 - *Ibid.* p. 921.
- 03 - *Ibid.* p. 922-923.
- 04 - PRICE, Derek J. de Solla. *O desenvolvimento da ciência: análise histórica, filosófica, sociológica e econômica.* Rio de Janeiro: Livros Técnicos e Científicos, 1976. 77 p. p.5-7.
- 05 - SARACEVIC, Tefko. *Information science: origin, evolution and relations.* Pre-print, 1991. 16 p.
- 06 - GRIFFITH, Delver C., SERVI, Patricia N., ANKER, Anita L. et al. The aging of scientific literature: a citation analysis. *Journal of Documentation*, v.35, n.3, p. 179-196, Sep. 1979.
- 07 - MATSUI, Sachiko, ASANO, Chooichiro, MATSUDA, Yoshiro. A system for generating publication statistics based on bibliographic information: bibliometric analysis for the development of economic and social thought. *Hitotsubashi Journal of Economics*, v. 30, n. 2, p. 121-156, Dec. 1989.

- 08 - BELL, Daniel. *O advento da sociedade pós-industrial: uma tentativa de previsão social*. São Paulo: Cultrix, 1977. p.139.
- 09 - RADA, Juan. Information technology and the Third World. In: FORESTER, Tom, ed. *The information technology revolution*. Cambridge, Mass.: MIT Press, 1985. p. 571-589.
- 10 - LUCAS, Clarinda Rodrigues. O papel do Sistema de Informação no processo de transferência de tecnologia. *Ciência da Informação*, v.16, n.2, p.181-191, jul./dez. 1987.
- 11 - CHRISTOVÃO, H. T. The aging of the literature of biomedical sciences in developed and developing countries. *Scientometrics*, v.7, p. 411-430, 1985.
- 12 - Cf. MATSUI, *op. cit.*



## 2. O PROCESSAMENTO TÉCNICO DE LIVROS E MONOGRAFIAS NOS SISTEMAS DE INFORMAÇÃO

Livros e monografias são documentos que tendem a abordar os assuntos de forma minuciosa, devido à suposta consolidação e maturação do conhecimento de seus autores ao longo do tempo, como também tendem a esgotar os assuntos que abrangem, sejam eles de natureza genérica ou específica. A redação desses documentos pressupõe uma pesquisa de campo profunda, uma revisão de literatura que cubra todos os tópicos tratados e, de certa forma, a criatividade do autor ao expor suas idéias. As citações incidem com maior intensidade em livros e monografias do que, por exemplo, nos artigos de periódicos.

Desta forma, ficam caracterizadas a extensão e a riqueza do conteúdo informativo desses documentos no processo de transmissão do saber.

Livros e monografias com qualidade de conteúdo veiculam informações específicas, de caráter teórico ou prático, que nem sempre tornam-se de domínio público, a não ser por aqueles que tomam ciência de sua edição, lêem-nos por inteiro e bem assimilam sua matéria.

Pode-se até supor que o conteúdo informativo específico de livros e monografias seja disseminado de maneira restrita, embora com reduzidíssima intensidade, somente através dos meios informais de comunicação, manifestação geralmente ocorrida entre cientistas e pesquisadores que, como já indicado, bem assimilam

seu conteúdo e, circunstancialmente, necessitam da informação consolidada no meio acadêmico para aplicação no embasamento teórico de seus trabalhos inovadores.

Um importante segmento de usuários da informação veiculada por livros e monografias são os estudantes de universidades, tanto a nível de graduação quanto de pós-graduação, os quais, sob a orientação de seus professores (geralmente também pesquisadores), recorrem a específicos tópicos de livros e monografias para a complementação de conhecimentos apreendidos ou para a formulação de pesquisas acadêmicas solicitadas. Os catálogos por assunto das bibliotecas universitárias geralmente permitem o acesso sub cabeçalhos ou termos de indexação genéricos.

Há também os usuários da informação jurídica. A literatura jurídica possui características muito peculiares, sendo a principal a expressiva veiculação de informações em livros e monografias, com uma abordagem bem específica dos assuntos tratados sempre vinculada à legislação e à jurisprudência afins.

Estes são alguns exemplos de usuários que recorrem intensamente à informação veiculada por livros e monografias e que, por sua vez, necessitariam dispor, nos Sistemas de Informação, de um acesso mais direto e imediato ao conteúdo informativo específico desses documentos.

Apesar de possuírem características informacionais diferentes dos demais documentos estabelecidos na comunicação formal do conhecimento, como artigos de periódicos e

"proceedings", livros e monografias têm seu conteúdo informativo pouco difundido pelos meios convencionais e informais de disseminação e recuperação da informação.

Verifica-se que as monografias e, principalmente, os livros, vêm recebendo por parte dos Sistemas de Informação um tratamento pouco condizente com o seu potencial informativo, com a atribuição de um número de cabeçalhos de assunto muito reduzido para posterior recuperação. A já mencionada inversão de valores ocorrida entre o livro e o periódico especializado certamente foi um dos fatores que contribuiu para essa questão (01). Esta é uma das grandes lacunas verificadas nos Sistemas de Informação quanto ao processamento técnico destes documentos.

### 2.1 - Catálogo e Política de Indexação

Na maioria dos Sistemas de Informação, a catalogação é o processo empregado no tratamento técnico das coleções de livros e monografias.

O processo de catalogação é integrado pela catalogação descritiva a qual, seguindo regras específicas (normalmente as estabelecidas pelo Anglo-American Cataloguing Rules - 2, editado pela American Library Association-ALA), objetiva descrever, de forma padronizada, as características bibliográficas dos documentos, e pela catalogação por assunto, procedimento normalmente orientado por uma Política de Indexação e por um vocabulário controlado, que objetiva identificar e atribuir

cabeçalhos de assunto que representem e expressem a sobrecedade do documento para posterior recuperação pelos usuários interessados. A catalogação por assunto pode ser efetuada empregando-se a pré ou a pós-coordenação de conceitos.

Entretanto, o processo de catalogação vem sendo desempenhado pelos Sistemas de Informação, de acordo com normas e padrões de procedimento que pouco evoluíram ao longo dos últimos anos, sobretudo a catalogação por assunto.

Essa estagnação refere-se, principalmente, à incompatibilidade dos objetivos do processo de catalogação com o desenvolvimento e o estágio atual da literatura contemporânea, assim como com a intensificação das necessidades de informação manifestadas pelos usuários dos Sistemas instituídos.

Através das Políticas de Indexação relativas a livros e a monografias normalmente vigentes nos Sistemas de Informação, percebe-se a instituição e o emprego de padrões informais quanto à quantificação e à qualidade da catalogação por assunto desse material bibliográfico. Quantidade pode não significar qualidade, e *vice versa*. Mas observa-se através do número de cabeçalhos de assunto comumente atribuídos por livro ou por monografia processada pelos Sistemas que a disponibilidade de acesso ao seu conteúdo informativo geral e específico fica sensivelmente comprometida.

Considerando sua extensão textual, seu propósito e a relativa rapidez de obsolescência de seu conteúdo informativo, os artigos de periódicos especializados, dada a sua importância

para os Sistemas de Informação, percebem um número de termos de indexação proporcionalmente bem superior ao que é atribuído aos livros e monografias (02).

Uma das razões comumente alegadas pelos Sistemas de Informação para a limitação do número de cabeçalhos de assunto a serem atribuídos por livro ou monografia processada é a sobrecarga de catálogos manuais ou mesmo de outros suportes físicos destinados ao armazenamento da representação temática dos documentos (espaço de memórias de computador, "diskettes" ou fitas magnéticas).

Também a dificuldade de interpretação rotineiramente verificada durante a análise conceitual de documentos é uma das razões para o comprometimento da qualidade da catalogação por assunto e para a atribuição de poucos cabeçalhos, os quais, em geral, possuem uma representatividade genérica ou então indevida. A propósito, é importante ressaltar que a atribuição de termos de indexação genéricos para a recuperação de informações oculta a verdadeira literatura existente sobre determinado assunto (03).

Como um exemplo bem apropriado, nas Ciências Humanas e Sociais verifica-se uma tendência entre pesquisadores e especialistas dedicados a determinado assunto de editarem, conjuntamente, alguns de seus trabalhos sob a forma de coletâneas (documentos, em inglês, denominados "key papers"), as quais veiculam informações sobre assuntos específicos relacionados a um tema mais abrangente. Neste caso, havendo limitações quantitativas, normalmente são atribuídos cabeçalhos

de assunto genéricos, ocorrendo uma perda considerável de informações supostamente relevantes para posteriores recuperações.

O melhor aproveitamento do conteúdo informativo de livros e de monografias a se obter através de uma catalogação por assunto mais exaustiva parece ser uma questão desconsiderada pelos Sistemas de Informação. Recuperações por autores e títulos são incompletas e insuficientes para o cumprimento dos objetivos fim dos Sistemas.

Na realidade, em relação a livros e a monografias, os Sistemas de Informação, durante o seu processamento técnico, conferem uma maior ênfase à catalogação descritiva (padrões de entrada de autores, pontuações, ordem de notas bibliográficas, ordem de pistas de fichas principais). No que se refere à catalogação por assunto, não se verifica uma tendência de se esgotar a representação de todo o potencial informativo destes documentos para posterior recuperação, quando então geralmente lhes é atribuído um reduzido número de cabeçalhos de assunto ou termos de indexação.

Não se percebe uma diretriz, ou uma preocupação coletiva ou mesmo individual, quanto à atual posição dos Sistemas de Informação no sentido de se alterar os procedimentos técnicos destinados a uma maior e proveitosa recuperação do conteúdo informativo de livros e monografias (04).

Outros fatores também devem concorrer para essa postura dos Sistemas de Informação, tais como: falta de recursos financeiros e humanos, limitações de capacidade operacional, Políticas de Indexação mal formuladas e, até mesmo e na pior das hipóteses, desinteresse profissional.

## 2.2 - Alterações necessárias no processamento técnico de livros e monografias

Cumprе reiterar o importante papel social e cultural das Bibliotecas e Centros de Documentação, os quais constituem-se um dos essenciais componentes do processo educacional e do desenvolvimento científico e tecnológico das nações. Investimentos de grande porte são efetuados pelas instituições mantenedoras para a sua organização, para a atualização de coleções bibliográficas e de procedimentos técnicos / operacionais, e manutenção.

No Brasil, a maior parte das Bibliotecas e Centros de Documentação/Informação possui grandes restrições de ordem financeira para atualizar e aprimorar seus acervos bibliográficos, principalmente o de livros e monografias por serem eles proporcionalmente mais dispendiosos e muito solicitados em termos de circulação. Assim sendo, revela-se como essencial que livros e monografias já incorporados ou a incorporar em um determinado acervo bibliográfico passem a ter seu conteúdo informativo mais explorado para futuras recuperações, o qual, na maioria dos casos, prescreve em médio

ou longo prazos, e que seu uso seja maximizado mediante uma maior acessibilidade.

Portanto, faz-se necessário alterar o atual posicionamento técnico dos Sistemas de Informação em relação a livros e monografias, ressaltando-se a qualidade e a importância das informações veiculadas por esses documentos no contexto dos Sistemas, através de um maior detalhamento da representação de seu conteúdo informativo.

Para se reverter a atual conduta dos Sistemas de Informação no sentido de pluralizarem a disponibilidade de acesso ao conteúdo informativo de livros e monografias, seria ideal ocorrer uma mudança radical na filosofia e na concepção do processo de catalogação por assunto ou de indexação, sensíveis alterações nas Políticas de Indexação ainda vigentes, como também um aprimoramento qualitativo da análise conceitual desses documentos, compatibilizando-a com os níveis de exaustividade ou de especificidade da catalogação por assunto ou da indexação a executar.

O conseqüente aumento dos custos operacionais dos Sistemas de Informação certamente seria compensado pelos benefícios advindos de uma expansão do dimensionamento de informações indexadas, permitindo-se, assim, um melhor atendimento aos usuários e, reiterando, a maximização de seu uso.

Paralelamente a uma proposta de alteração na concepção e no *modus operandi* da catalogação por assunto, que vise maximizar o uso de documentos tão importantes como os livros e as



monografias, faz-se necessário apresentar e sugerir à Comunidade da Informação mecanismos capazes de conferir, concomitantemente, aos profissionais dedicados a esse trabalho, objetividade, facilidade e precisão durante os processos de análise conceitual, de seleção de palavras com elevada representatividade conceitual e de atribuição de cabeçalhos de assunto em número proporcionalmente condizente com a especificidade da abordagem que os autores conferiram aos assuntos versados em seus trabalhos.

A fonte com maior probabilidade de precisão para a identificação de palavras e/ou expressões potencialmente representativas de informações difundidas por livros e monografias é a sua própria estrutural textual.

Há muito, e por razões óbvias, que os profissionais atuantes em Bibliotecas e Centros de Documentação utilizam os textos dos documentos como fontes indicativas de sua sobrecedência para o processo de indexação por assunto (05). Entretanto, este procedimento praticamente não é estudado, discutido e avaliado por pesquisadores e pelos próprios profissionais atuantes, e, por conseqüência, ainda não foi convertido em normas que o estabelecesse, perante a Comunidade da Informação, como um método capaz de conferir padrões ou modelos para aplicação e uso coletivos (06).

## CITAÇÕES E NOTAS

- 01 - Após exaustiva pesquisa bibliográfica, foi constatado que não existe um estudo especificamente dedicado a uma avaliação quantitativa e/ou qualitativa da atribuição de cabeçalhos de assunto ou de termos de indexação para livros e monografias. Entretanto, não é leviano afirmar que, nos Sistema de Informação, esses documentos percebem um reduzidíssimo número de pontos de acesso para posterior recuperação em proporção à riqueza e à extensão discursiva de seu conteúdo informativo. Por experiência acadêmica e profissional da autora, quanto às diretrizes e aos padrões de procedimentos técnicos relativos a livros e monografias verificados em Bibliotecas e Centros de Documentação/Informação, pode-se inferir que, nessas Instituições, há uma "cultura" de caráter profissional, a qual tendenciona a sua representação conceitual genérica e pouco expressiva, ocasionando grandes restrições de acesso a seu conteúdo.
- 02 - VICKERY, D. C. The Process of subject indexing. In: \_\_\_\_\_  
*Techniques of information retrieval*. London:  
Butterworths, 1970. p. 76-86.
- 03 - POULSEN, Claus. Subject access to new subjects, specific paradigms and surveys: PARADOKS - registration.  
*Libri*, v.40, n. 3, p. 179-202, 1990.

04 - NEILL, S. D. The dilemma of the subjective in information organization and retrieval. *Journal of Documentation*, v. 43, n. 3, p. 193-211, Sept. 1987.

05 - VICKERY, *op. cit.*, p. 79-80.

06 - VICKERY, B. C. The description of documents. In: ..... *On retrieval system theory*. 2nd ed. London: Dutterworths, 1968. p. 20-32.

Com o intuito de se obter o embasamento necessário para a formulação da investigação ora apresentada, assim como identificar a existência de prováveis pesquisas similares, foi efetuado um levantamento bibliográfico relativo ao problema anteriormente apontado, o qual, no âmbito da Ciência da Informação, demonstrou e confirmou serem proporcionalmente raras as pesquisas dirigidas ao melhor aproveitamento do conteúdo informativo de livros e monografias.

A nível internacional, a contribuição precursora é oferecida pela Professora Pauline ATHERTON COCHRANE (01, 02, 03, 04 e 05), que dirigiu o Subject Access Project - SAP, uma pesquisa que no final da década de 70 procurou comprovar e ressaltar a necessidade de se aumentar os índices de recuperação "online" de informações contidas em livros, principalmente através de OPACs (Online Public Access Catalogs). O projeto foi patrocinado pelo Council on Library Resources dos Estados Unidos e conduzido na Syracuse University, NY, e uma das técnicas propostas por COCHRANE para se aumentar o acesso ao conteúdo informativo de monografias é a utilização de palavras ou expressões extraídas de seus sumários e índices.

A pesquisa consistiu no acréscimo de uma média de 30 termos de indexação idênticos extraídos dos sumários e índices aos registros bibliográficos de uma amostra de livros assentados em formato MARC (Machine Readable Cataloguing da Library of Congress, Estados Unidos). Foi então constituída uma nova base de dados, denominada BOOKS, a qual, em paralelo, gerou um

vocabulário formado pelas palavras ou expressões dos sumários e índices, e destinado a pesquisas "online".

Após a inclusão dos termos de indexação na base BOOKS, foi efetuado um teste de recuperação de informações em uma amostra de 90 buscas realizadas concomitantemente na base MARC e na base BOOKS. Os principais resultados comparativos foram:

Recuperação de documentos relevantes: MARC 56 e BOOKS 130; somente 14 documentos relevantes foram recuperados em ambas as bases;

índice de precisão: MARC 35% e BOOKS 46%, e

Tempo dispendido nas pesquisas: MARC apontou, em média, o dobro do tempo gasto na base BOOKS.

Estas informações foram extraídas do capítulo especialmente dedicado à descrição dos resultados do "Subject Access Project-SAP", onde Pauline ATHERTON COCHRANE também descreve as três fases definidas para a realização do Projeto e uma breve análise de custo-benefício. Entretanto, não há uma descrição clara acerca das regras estabelecidas para a seleção das palavras nos sumários e índices. Ainda nesse capítulo, foi apresentado um estudo exploratório compreendendo as bases MARC, BOOKS e PRECIS (Preserved Context Index System) da British Library (06). Aliás, a recuperação de informações em Sistemas em Rede é um dos campos de pesquisa ao qual Pauline ATHERTON COCHRANE também dedica grande interesse (07).

Na ocasião dos trabalhos do SAP, Pauline ATHERTON COCHRANE contou com a colaboração de outros pesquisadores, dentre eles Barbara SETTEL, Keith W. RUSSEL e Carol MANDEL que, por sua vez, estudaram e participaram da produção de contribuições para a Biblioteconomia e para a Ciência da Informação relativas às investigações que realizavam.

Barbara SETTEL (08) editou um "paper" contendo regras e procedimentos para se promover o aumento de descrições por assunto incluídas em catálogos manuais de bibliotecas, seguindo, supostamente, a metodologia empregada durante o SAP. Em 1982, a mesma pesquisadora publicou um artigo sobre o mesmo tema mas já agora voltado para os catálogos "online" (09).

Keith W. RUSSEL (10) elaborou um relatório sobre os trabalhos apresentados e as conclusões de um encontro promovido pelo Council on Library Resources em Dublin, OH, 1982.

Mesmo através de insistentes correspondências remetidas a esses dois pesquisadores, não foi possível se obter uma cópia ou mesmo um resumo do conteúdo de seus principais trabalhos (08 e 10) para se efetuar uma descrição do método por eles aplicado e de seu desenvolvimento, assim como de se avaliar os resultados.

Segundo informações obtidas junto à Professora Pauline ATHERTON COCHRANE (11), o "Subject Access Project - SAP" não teve continuidade, e nunca foi adotado ou sequer analisado por instituições americanas como a Library of Congress ou por outras bibliotecas usuárias ou não do formato MARC.

As principais propostas do SAP foram posteriormente reestudadas e até mesmo empregadas em trabalhos independentemente desenvolvidos por outros pesquisadores interessados.

Dentre os principais, inclui-se as contribuições de:

a) Carol A. MANDEL (12 e 13) que publicou artigos dedicados ao enriquecimento de catálogos "online" para recuperações por assunto, seguindo a metodologia proposta pelo SAP.

b) Na Suécia, a metodologia do SAP foi empregada para a indexação de relatórios governamentais, legislação, decisões judiciais e normas administrativas. Esse trabalho foi executado sob a coordenação do Professor Bjorn V. TELL que contou com a colaboração de Irene WORMELL e vem sendo utilizado já há alguns anos, possuindo uma atualização "online" constante feita pelo DAFA - National Computing Centre for Administrative Data Processing, que gerencia a base de dados constituída. Segundo o Professor Bjorn V. TELL (14), palavras e expressões contidas nos títulos das seções de sumários que compreendam 5 ou mais páginas de texto são utilizadas para a indexação dos documentos, sem qualquer controle de linguagem de indexação. Continuando, o Professor menciona que "... a entrada de dados é feita por sua própria Secretária, por ser um processo de fácil execução, não requerendo um profissional especificamente treinado para a função, como também considera menos dispendioso indexar livremente os documentos, do que se recorrer a um vocabulário controlado. Ele ressalta que a terminologia jurídica empregada

nos documentos oficiais indexados é bem apropriada para uma indexação livre ..." (14). Além de sua participação efetiva, Bjorn V. TELL difundiu em Simpósios e Conferências internacionais as inovações e técnicas do trabalho que desenvolveu com a metodologia do SAP (15 e 16).

Irene WORMELL dedicou sua Tese de Doutorado ao SAP, sob a orientação do Professor Bjorn V. TELL, na Lund University, Suécia, em 1985 (17). Desde a época das pesquisas do SAP na Syracuse University, Irene WORMELL vem investigando e empregando processos que promovam o aumento do acesso ao conteúdo informativo de monografias e mantém contatos frequentes com a Professora Pauline ATHERTON COCHRANE nos Estados Unidos (09). Infelizmente, somente foi possível ter ciência do trabalho de Irene WORMELL através de seus artigos publicados em periódicos especializados (18, 19 e 20). O Professor Bjorn V. TELL gentilmente forneceu o endereço atual de Irene WORMELL mas a pesquisadora não respondeu às insistentes correspondências que lhe foram remetidas. Por essa razão, não foi possível obter maiores detalhes acerca de sua Tese de Doutorado, principalmente os resultados alcançados durante a investigação.

c) Outra pesquisadora, Mary MICCO, em 1980 defendeu sua Tese de Doutorado através de um estudo exploratório e comparativo de três conhecidas linguagens de indexação disponíveis para emprego nas Ciências Médicas e afins: LCSH (Library of Congress Subject Headings), MeSH (Medlars Subject Headings) e PRECIS (Preserved Context Index System) da British Library (21). A partir desse trabalho, Mary MICCO continuou realizando estudos e pesquisas referentes à



análise de conteúdo (22 e 23). Mary MICCO participou da experiência de aplicação da metodologia do SAP realizada na Austrália, na Biblioteca da Australian Defense Force Academy ADFA, onde uma média de 21 termos de indexação extraídos de sumários de livros foram adicionados aos registros de 6.000 livros incluídos em um Sistema de formato MARC. Os resultados do estudo indicaram que as buscas nesses livros com maior disponibilidade de acesso "online" por assunto alcançaram índices de revocação bem superiores às buscas efetuadas por títulos ou pelos cabeçalhos de assunto do vocabulário controlado (24).

Uma melhor descrição do método e dos resultados da experiência australiana foi relatada por Mary MICCO e Linda RAMBLER em estudo recentemente publicado (25). O protótipo do Sistema experimentalmente constituído compreendeu os registros já indexados com palavras ou expressões extraídas dos sumários durante o processo de catalogação por assunto. O *Thesaurus* "online" era sempre atualizado à medida que novos termos eram inseridos na base de dados. A principal inovação do protótipo consistiu em um "browsing" dos sumários dos livros selecionados, ou seja, havia "online" uma reprodução dos sumários que objetivava familiarizar os usuários com a coleção da Biblioteca.

Os resultados da experiência, a qual visou principalmente promover o aumento do uso da coleção de livros da Biblioteca, demonstraram que somente a adição de termos de indexação extraídos dos sumários provocaram um acréscimo de 300% na revocação, com aceitáveis índices de falsas recuperações.

Entretanto, é importante ressaltar que as autoras não descreveram os critérios qualitativos e quantitativos para a extração das palavras dos sumários dos livros destinadas à sua indexação.

Independente da descontinuidade do Subject Access Project - SAP nos Estados Unidos, essas experiências nele baseadas fazem supor uma provável exequibilidade de sua metodologia no sentido de se aumentar o acesso ao conteúdo informativo de livros e monografias.

Além das pesquisas relacionadas com o SAP, outros trabalhos que indicassem a promoção do acréscimo de pontos de acesso ao conteúdo informativo de livros e monografias foram identificados.

Jessica L. MILSTEAD publicou um livro especialmente direcionado para a recuperação da informação por assunto (26), onde, além de mencionar e sugerir fontes de palavras para a indexação, enfatiza a necessidade de se promover o aumento das possibilidades de acesso ao conteúdo temático de livros e monografias, especialmente os livros. Jessica L. MILSTEAD, cita o trabalho de Pauline ATHERTON COCHRANE junto ao SAP, afirmando que a possibilidade de maximização do uso de coleções de monografias foi significativamente aumentada com o emprego da metodologia utilizada no "Subject Access Project-SAP" (27). Cabe informar que o livro de Jessica L. MILSTEAD é enfaticamente citado por Mary MICCO no artigo que publicou, intitulado "Knowledge representation" (22).

Virgil DIDDATO realizou uma pesquisa sobre o acréscimo de termos de indexação originados dos sumários e índices de 125 livros aos cabeçalhos de assunto da LCSH atribuídos durante a catalogação desses documentos, com o objetivo de determinar a proximidade da terminologia empregada na representação conceitual (termos e cabeçalhos) com a terminologia utilizada por usuários para a descrição temática dos mesmos livros. Os resultados indicaram que o LCSH alcançou uma proximidade de 72,4%, os termos extraídos dos sumários e índices 81,3% e, empregando-se essas duas formas de representação, o índice chegou a 97,3%. Virgil DIDDATO também menciona a importância do SAP diante dos acréscimos de acesso por assunto que o Projeto conferiu aos registros MARC empregados na pesquisa (28).

George EICHINSKI ressalta a importância e a necessidade de se enriquecer os catálogos "online" das Bibliotecas, com uma maior atribuição de cabeçalhos de assunto por monografia processada (como também e principalmente livros) e reporta-se às experiências do SAP nos Estados Unidos e da ADFa na Austrália. Na ocasião, George EICHINSKI menciona o pequeno interesse dos "catalogadores" em efetivamente discutir o enriquecimento dos registros bibliográficos manifestado durante a "Conference for AARL" em 1981, à qual compareceu, o que permite supor-se a falta do estabelecimento de um pensamento coletivo e de uma mobilização profissional para se alterar a conduta dos Sistemas de Informação no que se refere à maior acessibilidade do conteúdo informativo de livros e de monografias (29).

Como trabalhos mais recentemente publicados, inclui-se o de Florence E. DEHART e Richard REITSMA que analisa a terminologia contida em sumários em relação a outras formas de representação conceitual, como os cabeçalhos de assunto da Library of Congress, e reflete uma posição restritiva dos autores quanto à atribuição de um número maior de cabeçalhos de assunto pela Library of Congress, sugerindo uma reavaliação do problema sob o ponto de vista sistêmico do MARC e dos "On line Public Access Catalogs-OPACs" (30).

Há também o trabalho de Clauss POULSEN o qual relata a experiência realizada na National Library of Education da Dinamarca para se aumentar o acesso ao conteúdo informativo de livros e monografias, usando-se uma variante da metodologia do "Subject Access Project-SAP". O sistema criado, denominado "PARADOKS - Registration", cobriu, inicialmente, o campo relativo aos meios de comunicação interativos aplicados à educação e, devido ao sucesso de seus resultados, foi estendido a todos os meios de comunicação utilizados na educação e ao ensino das ciências, da matemática e da tecnologia. Ainda nesse trabalho, o autor menciona uma maior demanda de acesso por assunto na literatura, ocasionada pela importância atribuída à informação pela sociedade atual, o que implicaria em uma melhoria da qualidade do acesso por assunto nos Sistemas de Informação, visto que os assuntos específicos são, de certa forma, "escondidos" sob termos de indexação mais abrangentes (31).

A maior parte dos trabalhos anteriormente citados nesta revisão da literatura encontram-se incluídos numa série constituída por três artigos sobre análise de conteúdo publicados na "Annual Review of Information Science and Technology-ARIST", em 1982, 1986 e 1989 (32, 33 e 34), principalmente os de Pauline ATHERTON COCHRANE, Mary MICCO, Virgil DIODATO, Carol MANDEL, Keith RUSSEL e Barbara SETTEL.

Pela descrição dos principais trabalhos devotados ao aumento da disponibilidade de acesso ao conteúdo informativo de livros e monografias, identificados a nível internacional, verifica-se que as pesquisas do SAP foram, além de pioneiras no campo, as que conferiram as bases metodológicas para aquelas posteriormente executadas. Entretanto, não foi possível apurar pela literatura e pelas correspondências pessoais mantidas com alguns dos pesquisadores citados quais as razões que ocasionaram a discretíssima repercussão deste trabalho na Comunidade de Informação norte-americana. Observa-se que as experiências efetivamente realizadas e acatadas ocorreram em outros países como: Suécia, Austrália e Dinamarca.

Finalizando, cumpre mencionar que, a nível nacional, apesar da realização de uma busca exaustiva, não foi identificada qualquer contribuição sobre o aumento de pontos de acesso ao conteúdo informativo de livros e monografias.

CITAÇÕES E NOTAS

- 01 - ATHERTON, Pauline. Improved subject access to books in on-line catalogs. In: INTERNATIONAL ON-LINE INFORMATION MEETING, 1st, London, 1977 [*Proceedings*]. Oxford: Learned Information, 1977. 234 p. p. 131-138.
- 02 - ATHERTON, Pauline et al. *Books are for use*: final report to the Council on Library Resources. Syracuse, NY: Syracuse University, School of Information Studies, 1978. 172 p.
- 03 - COCHRANE, Pauline A. Modern subject access in the online age. *American Libraries*, v. 15, p. 336-339, May 1984.
- 04 - \_\_\_\_\_. *Redesign of catalogs and indexes for improved online subject access*: selected papers of Pauline A. Cochrane. Phoenix: Dryx Press, 1985. 484 p.
- 05 - \_\_\_\_\_. The management of subject cataloguing in an era of interactive catalogues and bibliographic networks. *Cataloguing Australia*, v.11, n. 4, p. 2-8, Dec. 1985.
- 06 - Cf. COCHRANE, "*Redesign ...*", 1985, *op. cit.*, p.391-460.
- 07 - Cf. COCHRANE, 1984, *op.cit.*

- 08 - SETTEL, Barbara, ed. *Subject description of books: a manual of procedures for augmenting subject descriptions in library catalogs.* Syracuse, NY: Syracuse University, School of Information Studies, 1977.
- 09 - SETTEL, Barbara, COCHRANE, Pauline A. Augmenting subject description for books in online catalogs. *Database*, v.5, n.4, p. 29-37, Dec. 1982.
- 10 - RUSSEL, Keith W., comp. *Subject access: report of a meeting sponsored by the Council on Library Resources, 1982, June 7-9, Dublin, OH.* Washington: Council on Library Resources, Inc., 1982. 108 p.
- 11 - Correspondências pessoais e contatos telefônicos mantidos com a Professora Pauline Atherton Cochrane entre fevereiro e abril de 1990.
- 12 - MANDEL, Carol A., HERSCHMAN, Judith. Online subject access: enhancing the library catalog. *Journal of Academic Librarianship*, v.9, p. 140-155, July 1983.
- 13 - MANDEL, Carol A. Enriching the library catalog record for subject access. *Library Resources & Technical Services*, v.29, n.1, p.5-15, Jan./Mar. 1985.
- 14 - Correspondências pessoais mantidas com o Professor Bjorn V. Tell entre abril e maio de 1990.

- 15 - TELL, Bjorn V., WORMELL, Irene. Factual data retrieval from a data base constructed according to the SAP-technique: a project incorporating Government Committee reports. In: INTERNATIONAL RESEARCH FORUM IN INFORMATION SCIENCE, 4., Boras, Sweden, 1981. *Proceedings*. Lund: Lund University, 1981. p. 135-149.
- 16 - TELL, Bjorn V. Deep analyses of book contents: fact retrieval from Government Committee reports using the SAP-technique online. In: ESSEN SYMPOSIUM, Essen, 1981. [*Proceedings*]. Essen: Gesamthochschulbibliothek, 1982. 5 v. v.3: Future of library networks, p. 116-126.
- 17 - WORMELL, Irene. *Subject Access Project-SAP: improved subject retrieval for monographic publications*. Lund: Lund University, 1985. (Doc. Thesis)
- 18 - \_\_\_\_\_. SAP: a project for subject description of books. *Tidsk. Dokum.*, v.34, n.1, p. 3-6, 1978.
- 19 - \_\_\_\_\_. SAP: a new way to produce subject descriptions of books. *Journal of Information Science*, n.3, p. 39-43, 1981.
- 20 - INGWERSEN, Peter, WORMELL, Irene. Means to improve subject access and representation in modern information retrieval. *Libri*, v.33, n.2, p.94-119, June 1980.



- 21 - MICCO, Mary. *An exploratory study of three subject access systems in medicine: LCSH, MeSH, PRECIS.* Pittsburgh: University of Pittsburgh, 1980. (Doc. Thesis)
- 22 - MICCO, Mary, SMITH, Irma, HSIAO, Su-Ann et al. Knowledge representation: subject analyses. *Library Software Review*, p. 02-07, Apr. 1987.
- 23 - Correspondência pessoal mantida com a Professora Mary MICCO entre março e maio de 1990.
- 24 - BYRNE, Alex, MICCO, Mary. Improving OPAC subject access: the ADFA experiment. *College and Research Libraries*, v.49, n.5, p. 432-441, Sept. 1988.
- 25 - MICCO, Mary, RAMBLER, Linda. *Cognitive overload: an emerging problem hypertext for improved subject access: a proposed solution* [s.l.: s.n.] 1989. 8 p.
- 26 - MILSTAED, Jessica L. *Subject access systems: alternatives in design.* New York: Academic Press, 1984. 212 p.
- 27 - Cf. *ibid*, p. 96.
- 28 - DIODATO, Virgil. Table of contents and book indexes: how well do they match readers' descriptions of books? *Library Resources & Technical Services*, v.30, n.4, p.402-412, Oct./Dec. 1986.

- 29 - EICHINSKI, George. Enhancing subject data : talk or action. *Cataloguing Australia*, v.11, n.4, p.9-18, Dec. 1985.
- 30 - DEHART, Florence E., REITSMA, Richard. Subject searching and table of contents in single-work titles. *Technical Services Quarterly*, v.7, n.1, p. 33-51, 1969.
- 31 - POULSEN, Claus. Subject access to new subjects, specific paradigms and surveys: PARADOKS-registration. *Libri*, v.40, n.3, p. 179-202, 1990.
- 32 - TRAVIS, Irene L., FIDEL, Raya. Subject analysis. *Annual Review of Information Science & Technology*, v.17, p. 123-157, 1982.
- 33 - SCHWARTZ, Candy, EISENMANN, Laura Malin. Subject analysis. *Annual Review of Information Science & Technology*, v.21, p.37-61, 1986.
- 34 - LANCASTER, F. W., ELLIKER, Calvin, CONNELL, Tschera Harkness. Subject analysis. *Annual Review of Information Science & Technology*, v.24, p.35-84, 1989.

#### 4 · OBJETIVOS

O objetivo principal da presente Dissertação é investigar a possibilidade de se aumentar os níveis de exaustividade da indexação de livros e monografias em Sistemas de Informação pós e pré-coordenados, sem, entretanto, comprometer os níveis de especificidade a ocorrer na recuperação de informações.

A elevação desses níveis de exaustividade seria alcançada através da atribuição de um número maior de cabeçalhos de assunto durante o processamento técnico de livros e monografias, os quais seriam originados da própria estrutura textual destes documentos.

Assim sendo, e em termos mais específicos, os objetivos da presente pesquisa são:

- Definir e propor um método de interpretação do conteúdo informativo de livros e monografias, que possa reduzir o esforço intelectual dos profissionais e, ao mesmo tempo, aprimorá-lo, baseado na identificação de conceitos relevantes representados pela terminologia empregada por seus autores;
- Identificar os elementos estruturais de livros e monografias que podem fornecer palavras e/ou expressões potencialmente representativas dos conceitos relevantes veiculados por esses documentos;

Demonstrar o potencial desses elementos estruturais para a identificação e seleção de palavras com expressiva representatividade conceitual, avaliando a aplicação da metodologia empregada pela investigação em Sistemas de Informação pós-coordenados e pré-coordenados, e

Avaliar a capacidade de representação de assuntos das linguagens de indexação pré e pós-coordenadas utilizadas pelos Sistemas de Informação a examinar, investigando, também, a sua proximidade com a linguagem natural dos autores de livros e monografias, através do emprego da terminologia corrente para a representação conceitual.

A consecução dos objetivos acima descritos poderá oferecer à Comunidade da Informação instrumentos teóricos e praticamente exequíveis, e destinados ao aprimoramento dos serviços prestados pelos Sistemas de Informação, no que se refere ao aumento da acessibilidade ao conteúdo informativo de livros e monografias e à maximização do uso das coleções destes documentos.

## 5 - JUSTIFICATIVA

Diante dos problemas apontados e dos objetivos a atingir anteriormente descritos, como também pelo fato da revisão de literatura não ter indicado um número expressivo de trabalhos, sobretudo a nível nacional, a presente investigação reveste-se de grande importância para o desenvolvimento e aperfeiçoamento dos processos técnicos dos Sistemas de Informação, como também os conseqüentes benefícios que esses processos proporcionariam para a sociedade, por permitirem uma maior e melhor acessibilidade ao conteúdo informativo de documentos tão relevantes quanto os livros e as monografias.

Quanto aos processos técnicos, a maior importância refere-se à proposição de mecanismos capazes de concorrerem para uma mudança radical no atual tratamento conferido aos livros e às monografias pelos Sistemas de Informação, compatibilizando-o com o real valor informativo destes documentos.

Para tanto, a metodologia aplicada pela investigação intenciona demonstrar a possibilidade de se promover a atribuição de um maior número de cabecalhos de assunto a serem identificados e extraídos dos elementos estruturais dos próprios documentos, com uma flexibilidade capaz de permitir indexações genéricas ou específicas, como também os desejáveis níveis de revocação e de precisão a atingir durante a recuperação de informações.

Este procedimento objetiva, também e principalmente, facilitar o trabalho de interpretação do conteúdo informativo do documento de forma mais próxima às idéias dos autores, na tentativa de se reduzir a inerente subjetividade do processo e o esforço intelectual do profissional que o executa, utilizando-se a terminologia corrente empregada no documento para a sua representação conceitual.

Um outro enfoque importante a ser abordado diz respeito à necessidade de se rever e reformular as diretrizes ou Políticas de Indexação relativas a livros e monografias normalmente vigentes nos Sistemas de Informação, sobretudo quanto às limitações quantitativas de atribuição de cabeçalhos, as quais, dentre outros pontos desfavoráveis, obrigam os profissionais a priorizar cabeçalhos, deixando, por vezes, de atribuir outros de grande relevância para o contexto do Sistema de Informação.

No Brasil, existe um bom número de Sistemas de Informação que dispõem de recursos automatizados para o assentamento e recuperação de registros, com grande capacidade de armazenamento de informações em memória, fato que permitiria a inclusão de um maior número de pontos de acesso ao conteúdo informativo de livros e monografias. Também para os Sistemas de Informação que ainda operam convencionalmente, mesmo com o aumento da sobrecarga de trabalho e de despesas administrativas, seria extremamente favorável intensificar o acesso e a utilização de coleções de livros e monografias por seus usuários.

A pesquisa também fornecerá subsídios suficientes para a realização de estudos posteriores quanto à análise de custo-benefício da sua aplicabilidade em Sistemas de Informação, que seriam observados e mensurados mediante a realização de um teste de recuperação de informações em coleções de livros e monografias indexados de acordo com a metodologia utilizada, durante determinado período.

Quanto aos benefícios para a sociedade advindos do aprimoramento dos serviços prestados pelos Sistemas de Informação, a presente investigação ofereceria uma contribuição ímpar, já que a maior disponibilidade de acesso ao conhecimento, em todos os seus níveis, bem como a sua assimilação, são fatores essenciais para a reversão, a médio e longo prazos, dos quadros social, científico, tecnológico e cultural de países como o Brasil.

Além disso, no cenário internacional, verifica-se que a transferência do conhecimento entre as nações tornou-se um forte quesito de exercício de poder e de dominação, como também hoje a informação é considerada um insumo econômico de altíssimo custo e indispensável para a inserção e participação das nações no âmbito das relações internacionais, sejam políticas, econômicas ou sociais.

Nos países integrantes do Terceiro Mundo, "... o homem não é conhecedor, mas usuário do conhecimento, ou pior, é usuário dos bens que incorporam em si o conhecimento ..." (01). Esta citação configura perfeitamente o quadro crítico das condições educacionais de suas populações e o seu subdesenvolvimento econômico, tecnológico e social, o que vem a restringir, sobremaneira, a sua capacidade de penetração e de respeitabilidade na conjuntura mundial.

Diversos componentes integrariam as ações necessárias para a modificação de tão lamentável conjunto de circunstâncias. Sob esse aspecto, mesmo que com uma participação contextual relativamente pequena, caberia aos Sistemas de Informação desses países dinamizar o uso apropriado dos meios informacionais já disponíveis em seus acervos, os quais são geralmente constituídos sob grandes sacrifícios, ou seja, restrições de ordem financeira e pouca disponibilidade de recursos humanos e operacionais.

Especificamente quanto aos livros e às monografias integrantes desses acervos, a presente investigação tenciona oferecer ao campo da Ciência da Informação uma contribuição de caráter empírico que vem apresentar e propor mecanismos destinados à melhoria e ao aumento da acessibilidade das informações tão básicas e fundamentais veiculadas por esses documentos.



CITAÇÕES

01 - DARDOSA, Denis Borges. O avanço do feudalismo  
informacional. *Jornal do Brasil*, Rio de Janeiro,  
14 jul. 1991. Caderno Idéias, p. 10-11.

## 6. MATERIAL

Para atingir os objetivos propostos pela presente investigação, a constituição da amostra de pesquisa previu a inclusão de livros e monografias que integrassem concomitantemente o acervo bibliográfico de dois Sistemas de Informação: um que adotasse a pós-coordenação, como técnica de indexação e de recuperação da informação, e outro que adotasse a pré-coordenação. Isso porque pós e pré-coordenação são os únicos processos técnicos instituídos e disponíveis para a indexação e para a recuperação temática de informações textuais.

O Sistema de Informação seguidor da pós-coordenação escolhido foi o da Biblioteca do Banco Nacional de Desenvolvimento Econômico e Social - BNDES (01), e para a pré-coordenação, o da Rede BIBLIODATA (CALCO) (02 e 03), gerenciada pela Fundação Getúlio Vargas, ambos localizados na cidade do Rio de Janeiro - RJ.

### 6.1 - A identificação da amostra

Como etapa inicial da identificação, o campo do conhecimento escolhido para a definição da amostra de livros e monografias a examinar foi ECONOMIA, por ser o de maior compatibilidade com as funções fim do Banco Nacional de Desenvolvimento Econômico e Social - BNDES.

Em fevereiro de 1990, o catálogo topográfico das diversas coleções bibliográficas integrantes do acervo da Biblioteca do BNDES registrava a existência de um expressivo número de livros e monografias sobre o campo Economia. Pelo mesmo catálogo topográfico, foram separados os livros e monografias mais recentemente incorporados ao acervo, ou seja, até 31.12.89, desconsiderando-se idioma, autoria ou o tema econômico específico abordado.

A seguir, foi verificado se os mesmos livros e monografias integrantes do acervo da Biblioteca do BNDES haviam sido também incorporados à Rede BIBLIODATA (CALCO). Foi constatado que cerca de 100 (cem) documentos idênticos (observando-se, principalmente, suas edições) integravam o acervo de ambos os Sistemas de Informação. Estes 100 livros e monografias foram então separados como uma pré-amostra.

#### 6.2 - O tamanho da amostra

DAILEY (04) afirma que "... o tamanho correto de uma amostra depende da natureza do universo a investigar e do propósito da pesquisa ..." e que "... cerca de 30 casos parece ser o número mínimo indispensável em pesquisas nas quais a análise de dados estatísticos deve ser efetuada, embora algumas técnicas podem ser usadas com menos de 30 casos ...".

Com base nessa orientação, 30 dos 100 livros e monografias integrantes da pré-amostra foram selecionados aleatoriamente, através de calculadora Hewlett-Packard, modelo 11C.

Essa seleção resultou na composição definitiva da amostra para a realização da pesquisa, ou seja, 30 livros e monografias.

As principais características da amostra então composta são:

a) Categorias dos documentos:

19 livros;

7 coletâneas; e

4 teses.

b) Autoria:

23 livros e monografias até 3 autores; e

7 coletâneas.

c) Idioma dos textos:

19 livros e monografias até 3 autores em português;

3 livros e monografias até 3 autores em inglês;

1 livro em espanhol;

6 coletâneas em português; e

1 coletânea em inglês.

Aos livros e às monografias então selecionados foi conferida uma numeração seqüencial, de 1 a 30, objetivando a sistematização da apresentação da análise e da discussão dos resultados.

Visando uma identificação detalhada dos livros e monografias integrantes da amostra a investigar, as referências bibliográficas, os cabeçalhos de assunto atribuídos na catalogação oficial e os números de chamada ou de identificação recebidos em ambos os Sistemas de Informação examinados, correspondentes a cada um dos 30 documentos selecionados, encontram-se no Anexo 1. As informações referentes aos livros e monografias da Biblioteca do BNDES foram levantadas em fevereiro de 1990 e as referentes à Rede BIBLIODATA (CALCO) foram obtidas através de microfichas editadas pela Fundação Getúlio Vargas intituladas " Lista por autor/título das obras cadastradas no BIBLIODATA/CALCO ", período 10.09.79 a 24.08.89 .

A seguir é apresentada e descrita a metodologia utilizada na investigação.

## CITAÇÕES E NOTAS

01 - O Banco Nacional de Desenvolvimento Econômico e Social-BNDES é uma instituição financeira do Governo Federal, criada em 1952, e tem como objetivo principal financiar, a longo prazo, projetos de implantação e expansão de empresas atuantes na maior parte dos segmentos da economia nacional tais como indústria, infra-estrutura, serviços, e, dessa forma, promover o desenvolvimento e o fortalecimento da economia brasileira. A Biblioteca do BNDES, também criada em 1952 para atender às necessidades de informação decorrentes das atividades meio e fim da instituição, possui um acervo constituído basicamente por livros, monografias (folhetos, Dissertações, Teses e estudos econômicos setoriais), periódicos especializados, anais de congressos e seminários, cobrindo, de um modo geral, os seguintes campos do conhecimento: economia, finanças nacionais e internacionais, direito, administração, informática e tecnologia (principalmente a industrial). Na década de 70 foi iniciado o trabalho de automação dos registros bibliográficos, o qual vem sendo aperfeiçoado em proporção à demanda de serviços que a Biblioteca presta a seus usuários. O "software" foi criado e desenvolvido pelos técnicos do Departamento de Sistemas do BNDES, com a participação direta dos bibliotecários, e, devido à sua concepção atual, estabelece como norma de indexação e de recuperação das informações a pós-coordenação, permitindo, inclusive, o uso da intersecção de termos de indexação.

02 - A Rede BIBLIODATA (CALCO), gerenciada pela Fundação Getúlio Vargas (Rio de Janeiro), foi criada na década de 70 com base no Sistema MARC (Machine Readable Cataloguing), da Biblioteca do Congresso dos Estados Unidos. É integrada por 83 Bibliotecas e Centros de Documentação, e ainda 443 Bibliotecas depositárias ou setoriais, perfazendo um total de 526 coleções distintas de documentos em processamento, localizadas em diferentes pontos do país (dados de 09.03.92), e tem como objetivo básico fornecer infra-estrutura para os serviços das Bibliotecas e Centros de Documentação, de forma cooperativa, possibilitando a agilização no processamento técnico, e evitando a duplicação de serviços de catalogação (descritiva e por assunto) e de classificação, permitindo, assim, o aperfeiçoamento técnico, a normalização e a redução de custos operacionais. A Rede BIBLIODATA (CALCO) dispõe de um catálogo coletivo composto pelos registros bibliográficos de todas as Bibliotecas e Centros de Documentação integrantes, que em 09.03.92 somava 446.539 registros, sendo 428.895 correspondentes a livros e monografias e 17.764 a periódicos. Deste total de registros, 139.001 já foram utilizados em catalogações cooperativas. A Rede BIBLIODATA (CALCO) é, portanto, o maior e mais importante registro bibliográfico integrado existente no Brasil, uma vez que as Bibliotecas e Centros de Documentação que a compõem possuem acervos de valor incalculável, sob os aspectos informativo e cultural, e também no que se refere ao registro e resguardo da memória literária brasileira, uma vez que a Biblioteca Nacional é a unidade que possui o maior número de registros incluídos na

Rede. Além disso, a Rede BIBLIODATA (CALCO) é uma Base de Dados bibliográficos que abrange informações das mais diversas áreas do conhecimento, devido à diversidade dos acervos bibliográficos das Bibliotecas e Centros de Documentação que a integram. A Base de Dados pode ser acessada através da rede pública de telefonia, via RENPAC (EMBRATEL) ou pelo FGVCOM, programa de comunicação desenvolvido pela própria FGV. As informações podem ser recuperadas em diferentes níveis, com a formação de argumentos de busca, através de terminal de computador ou microcomputador emulando terminal.

- 03 - FUNDAÇÃO GETÚLIO VARGAS. Centro de Processamento de Dados. Gerência de Sistemas de Acesso Público. *Calco on-line: manual do usuário* [Rio de Janeiro: 1989?] 22 f.
- 04 - DAILEY, Kenneth D. *Methods of social research*. New York: Free Press, 1978. p. 84.



## 7. METODOLOGIA

A metodologia empregada na investigação foi constituída por quatro etapas efetuadas na sequência a seguir descrita:

- a) Definição dos elementos estruturais de livros e monografias a utilizar na análise conceitual;
- b) Análise conceitual;
- c) Compatibilização dos cabeçalhos de assunto formados a partir da análise conceitual com os vocabulários controlados adotados pelos Sistemas de Informação investigados; e
- d) Estudo da distribuição, por páginas de texto, dos níveis de subordinação dos sumários dos documentos integrantes da amostra.

### 7.1 Definição dos elementos estruturais de livros e monografias a utilizar na análise conceitual

Os principais elementos estruturais de livros e monografias que podem atuar como fontes de identificação de palavras e/ou expressões passíveis de uso no processo de indexação por assunto são: título, subtítulo (quando houver), sumários, prefácios, apresentações, o texto propriamente dito, referências bibliográficas e índices (01 e 02).

Os prefácios e apresentações, nem sempre redigidos pelos autores dos documentos, são elementos estruturais meramente formais, cujos textos são, por vezes, desvincilhados do seu propósito informativo.

Recorrer ao texto completo de livros e monografias não é recomendado, dada a sua extensão, embora seja operacional e tecnicamente viável, haja visto que a própria indexação automática é indicada para a indexação de documentos com textos de menor dimensão (artigos de periódicos e resumos, por exemplo).

As referências bibliográficas e citações são uma fonte de identificação de termos de indexação, através dos títulos dos documentos que descrevem, não somente pelas relações temáticas, como também por demonstrarem a contextualização que o autor concedeu ao livro ou à monografia. Mesmo assim, as palavras e/ou expressões usadas nos títulos dos documentos citados ou consultados podem não representar, de modo ideal, os conceitos mais relevantes incluídos em livros e monografias, uma vez que, eventual ou naturalmente, não correspondem à especificidade dos assuntos percorridos em sua estrutura textual. A literatura derivativa (referências bibliográficas e citações) incluída em um livro ou em uma monografia, em geral, não pode ser considerada como homogênea em relação à especificidade da abordagem destes documentos, pois é praxe os autores recorrerem a documentos de, a documentos sobre e a documentos correlatos para obterem os embasamentos teóricos e empíricos necessários à elaboração de seus trabalhos (03).

Os índices, em sua maioria, não são elaborados pelos autores de livros e monografias, e suas entradas temáticas e/ou onomásticas, por resultarem de um processo interpretativo, podem representar conceitos de forma não muito satisfatória. Outro ponto desfavorável para os índices: as palavras, expressões e nomes próprios que os integram são ordenados alfabeticamente, o que, de certa maneira, não permite a identificação ideal de relações conceituais.

Os títulos, subtítulos e sumários são fontes riquíssimas de palavras e/ou expressões representativas da sobreidade de documentos, principalmente de livros e monografias:

a) Títulos e subtítulos:

- . são redigidos pelos autores;
- . expressam a contextualização mais ampla dos documentos;
- . fonte comumente utilizada pelos profissionais da informação durante a indexação; e
- . estudos e experiências dedicados à indexação de documentos através dos títulos indicam que eles são uma proveitosa fonte de palavras e/ou expressões representativas de conteúdo informativo, embora insuficientes e inadequadas para uma indexação por assunto completa, já que expressam a contextualização mais geral do documento a que correspondem.

b) Sumárias:

os títulos de suas seções ou partes são também redigidos pelos autores;

expressam toda a contextualização do documento, uma vez que apresentam, estruturadamente, a abordagem que o autor concedeu ao seu trabalho, da mais geral à mais específica, podendo, portanto, fornecer indicações completas e precisas quanto aos conceitos veiculados pelo livro ou pela monografia; e

a estrutura dos sumários configura perfeitamente os diferentes níveis de especificidade da abordagem atribuída pelos autores, permitindo ao indexador utilizar as palavras, expressões e nomes próprios de acordo com a Política de Indexação vigente no Sistema de Informação que trabalha.

Pelo exposto, verifica-se que as restrições da utilização de títulos na indexação por assunto de livros e monografias podem ser superadas, recorrendo-se, também e concomitantemente, aos sumários como fonte de palavras e expressões para a representação de conceitos.

A principal e mais importante justificativa de se utilizar conjuntamente títulos, subtítulos e sumários na catalogação por assunto de livros e monografias, além das já apresentadas, é que por serem elementos estruturais redigidos pelo autor, suas

palavras e/ou expressões são, evidentemente, aquelas mais próximas de suas idéias e, portanto, potencialmente utilizáveis para a representação apropriada de conceitos.

Em alguns casos, os sumários podem incluir em sua estrutura palavras e/ou expressões com uma representação concisa e, portanto, não muito clara quanto ao conteúdo semântico, a qual, para fins de indexação, deve ser complementada recorrendo-se à leitura do texto do documento correspondente (04).

Ainda assim, os títulos, subtítulos e sumários, conjuntamente, parecem ser as melhores fontes da estrutura de livros e monografias para a identificação de palavras e/ou expressões que bem representem o seu conteúdo informativo, sendo, portanto, os elementos estruturais utilizados na presente investigação.

#### 7.1.1 - Os títulos, subtítulos e sumários.

Com a opção pelos títulos, subtítulos e sumários como elementos estruturais a utilizar como fontes de palavras e/ou expressões para a representação de conceitos, e objetivando apresentar o material utilizado durante a presente investigação, as cópias completas dos sumários pertencentes aos 30 livros e monografias integrantes da amostra e as respectivas referências bibliográficas compõem o Anexo 2, no qual encontram-se dispostos na seqüência da numeração que receberam (ver item 6.2).

Através das cópias dos sumários, pode-se observar a heterogeneidade da amostra, uma vez que os sumários apresentam estruturas e níveis de especificidade bem distintos.

Com o objetivo de melhor precisar os resultados da pesquisa, no sentido de apontá-los de forma coerente com as características da amostra, e, ao mesmo tempo, respeitando-se a estrutura hierárquica de abordagem dos assuntos atribuída pelos autores à contextualização de seus trabalhos, foram conferidos a cada um dos sumários os níveis de subordinação máximos identificados em suas estruturas.

Estes níveis de especificidade ou de subordinação de assuntos foram identificados e definidos através da apresentação gráfica e da própria estrutura dos sumários, considerando-se para tanto: numeração das partes dos sumários (numeração progressiva e/ou numeração com algarismos romanos para Capítulos e Partes), indentação gráfica, ou então, subordinação visual permitida pelo uso de caracteres gráficos distintos. Em caso de uso concomitante de mais de uma técnica de numeração ou de apresentação gráfica, foi respeitada a estrutura conferida aos sumários pelos autores.

Já que os elementos estruturais dos livros e monografias utilizados para a execução da pesquisa foram os títulos, subtítulos e sumários, os níveis máximos de subordinação de abordagem de assuntos ficaram assim estabelecidos:

- para os sumários: níveis de subordinação de 1 até 5; e
- para os títulos e subtítulos: nível de subordinação 0.

Esta distinção de níveis de subordinação entre os títulos, subtítulos e sumários foi proposital, no sentido de se poder verificar os resultados da pesquisa isoladamente, ou seja: resultados obtidos através dos títulos e subtítulos, e resultados obtidos na seqüência dos níveis de subordinação dos sumários.

O Anexo 3 apresenta os níveis máximos de subordinação atribuídos aos sumários dos 30 livros e monografias integrantes da amostra.

As diferenças estruturais dos sumários provocaram a divisão da amostra em grupos de documentos com estruturas de sumários semelhantes, de acordo com o nível máximo de subordinação de seus sumários.

Por essa razão e em função dos níveis de subordinação definidos, foram constituídos 5 Grupos de livros e monografias na amostra investigada, assim especificados:

- Grupo 1: integrado por livros e monografias cujos sumários possuíam estrutura até o nível 1 de subordinação; este grupo incluiu os documentos de número 27, 28, 29 e 30;

b) Grupo 2: integrado por livros e monografias cujos sumários possuíam estrutura até o nível 2 de subordinação; este grupo incluiu os documentos de número 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 e 26;

c) Grupo 3: integrado por livros e monografias cujos sumários possuíam estrutura até o nível 3 de subordinação; este grupo incluiu os documentos de número 5, 6, 7, 8, 9 e 10;

d) Grupo 4: integrado por livros e monografias cujos sumários possuíam estrutura até o nível 4 de subordinação; este grupo incluiu os documentos de número 3 e 4; e

e) Grupo 5: integrado por livros e monografias cujos sumários possuíam estrutura até o nível 5 de subordinação; este grupo incluiu os documentos de número 1 e 2.

Desta forma, a investigação foi efetuada nos seguintes Grupos constituídos para os 30 livros e monografias selecionados:

- 4 documentos incluindo até o nível 1 de subordinação;
- 16 documentos incluindo até o nível 2 de subordinação;
- 6 documentos incluindo até o nível 3 de subordinação;
- 2 documentos incluindo até o nível 4 de subordinação; e
- 2 documentos incluindo até o nível 5 de subordinação.



O nível 0 de subordinação (títulos e subtítulos) foi examinado em todos os Grupos.

Cabe ressaltar que, da estrutura dos sumários, foram desconsiderados como tópicos de identificação de palavras e/ou expressões representativas de conteúdo: apresentações, prefácios, introduções, conclusões, tabelas, anexos, listas de quadros, apêndices, ou então tópicos com denominações e características semelhantes.

## 7.2 - Análise conceitual

Análise conceitual, ou análise de conteúdo como também é denominada, é a identificação de conceitos veiculados em um documento, expressos por palavras ou conjuntos de palavras, que vêm ao encontro dos objetivos, interesses e usuários de um Sistema de Informação.

Os conceitos identificados como relevantes para um Sistema de Informação são, portanto, os elementos básicos e iniciais dos processos de catalogação e de indexação por assunto, os quais, de acordo com a Política de Indexação estabelecida para o Sistema, deverão ser posteriormente compatibilizados com uma linguagem de indexação, para se tornarem pontos de acesso ao conteúdo informativo do documento.

A análise conceitual é um processo altamente subjetivo porque implica em interpretações pessoais das idéias do autor para a seleção de conceitos pertinentes e, portanto, requer do indexador experiência, domínio da terminologia do campo do conhecimento abrangido pelo documento e, concomitantemente, a observância do contexto do Sistema de Informação detentor do documento (objetivos fim, usuários e capacidade operacional, por exemplo).

Uma vez que foi previsto executar a investigação em livros e monografias idênticos integrantes dos acervos de dois Sistemas de Informação, um adepto da pós-coordenação e outro adepto da pré-coordenação, foi definido que a análise conceitual a ser efetuada através das palavras e/ou expressões incluídas nos títulos, subtítulos e sumários, seria dirigida exclusivamente à Biblioteca do BNDES, devido à familiaridade da autora para com o Sistema de Informação e com a terminologia do campo Economia, apesar da distinção existente entre os contextos dos dois Sistemas investigados.

Para a Rede BIBLIODATA (CALCO) não foi realizada uma análise conceitual específica, uma vez que a amostra da investigação era composta por livros e monografias incluídos simultaneamente no acervo de algumas das Bibliotecas e Centros de Documentação integrantes da Rede, os quais, certamente, possuíam objetivos e usuários com características bem diferentes. Além disso, seria metodologicamente inviável e também desnecessário para a comprovação dos resultados junto à Rede BIBLIODATA (CALCO) executar uma análise conceitual

exclusiva.

É importante esclarecer e frisar que os conceitos a se identificar durante a análise conceitual dirigida à Biblioteca do BNDES seriam os mesmos a utilizar para a realização da pesquisa junto à Rede BIBLIODATA (CALCO).

A análise conceitual foi executada através dos títulos, subtítulos e dos tópicos dos sumários integrantes da amostra, observando-se os níveis de subordinação atribuídos aos elementos acima mencionados, definidos como de 0 a 5.

A análise conceitual objetivou a identificação e a seleção de palavras e/ou expressões representativas do conteúdo informativo dos livros e monografias e foi efetuada obedecendo a seqüência a seguir descrita.

#### 7.2.1 - Seleção de palavras e/ou expressões

Para a identificação e seleção das palavras e/ou expressões que pudessem representar conceitos semanticamente úteis para a Biblioteca do BNDES, foram consultadas algumas contribuições já consolidadas na literatura sobre a análise de conteúdo.

De acordo com sua concepção taxionômica, RANGANATHAN, ao elaborar a "Colon Classification" (OS), criou um Sistema de Classificação analítico-sintético, no qual todo ou qualquer parte do conhecimento humano pode ser representado de acordo com características e relações pertinentes aos contextos geral ou

específico de determinado documento. Essas características e relações foram denominadas facetas e aquelas de nível de abordagem mais amplo incluídas na "Colon Classification", também chamadas de Entidades Básicas (Basic Entities) são: Personality, Matter, Energy, Space e Time (PEMESI). A concepção de RANGANATHAN permitiu profundas alterações teóricas e conceituais no estudo e na utilização dos Sistemas de Classificação, uma vez que ofereceu à Biblioteconomia e à Ciência da Informação um método flexível e contextual quanto à representação conceitual para fins de organização e recuperação da informação.

O "Classification Research Group" (CRG), entidade criada em 1952, e estabelecida em Londres, congrega alguns especialistas que se dedicam a estudos e pesquisas voltados para os Sistemas de Classificação Bibliográfica destinados à representação do acervo de Sistemas de Informação. Ao longo dos anos, o CRG tem sido uma entidade profundamente adepta da teoria contida na "Colon Classification", desenvolvendo e aprimorando os seus princípios, inclusive sugerindo aplicações na análise conceitual e na indexação de documentos.

VICKERY (06, 07 e 08), membro do CRG, a partir das concepções filosófica e teórica de RANGANATHAN relativas à classificação do conhecimento, propôs uma forma de caracterização da estrutura do conhecimento através da denominação e da definição de categorias semânticas ou facetas, ou seja, as particularidades pelas quais se considera um conceito. Essa abordagem foi fundamental para o desenvolvimento da metodologia do presente trabalho, uma vez que contribuiu para a definição do *modus operandi* da análise conceitual, por

orientar e permitir a contextualização dos conceitos representados pelas palavras e/ou expressões incluídas nos títulos, subtítulos e sumários dos documentos integrantes da amostra. VICKERY (cf.08 ) afirma que, no âmbito da ciência e da tecnologia, pode-se constatar a utilização profícua das categorias semânticas para a representação de conceitos.

LANGRIDGE (09), também integrante do CRG, e adepto da concepção sistêmica de RANGANATHAN para a classificação e representação do conhecimento, apresenta categorias semânticas muito semelhantes às definidas e denominadas por VICKERY.

As categorias semânticas estabelecidas por VICKERY e por LANGRIDGE apresentam um grau de flexibilidade ainda maior em relação às facetas definidas por RANGANATHAN, uma vez que podem orientar os profissionais e serem utilizadas na indexação de documentos referentes a diversos campos do conhecimento com maior facilidade que a *Colon Classification*.

As contribuições de RANGANATHAN, VICKERY e LANGRIDGE foram extremamente proveitosas, sob os aspectos teórico e prático, porque forneceram as bases para a análise conceitual efetuada durante a presente investigação, permitindo a formulação de um esquema identificador para a representação de conceitos. Tal esquema orientou todo o processo de análise conceitual e de seleção de palavras, expressões e nomes próprios realizado nos títulos, subtítulos e sumários dos documentos integrantes da amostra.

O esquema **identificador** de conceitos ficou assim constituído:

PARTICULARIDADES DOS CONCEITOS A REPRESENTAR (FACETAS)		EXEMPLOS DE REPRESENTAÇÃO DE CONCEITOS
ENTIDADES BÁSICAS (*)	CATEGORIAS SEMÂNTICAS	
	PESSOAS	MILTON FRIEDMAN
Personality	COISAS COMO UM TODO	ECONOMIA
	TIPOS DE COISAS	ECONOMIA KEYNESIANA
	PARTES DE COISAS	MACROECONOMIA, MICROECONOMIA
Matter	MATERIAIS	MOEDA, CAPITAL DE GIRO
	PROPRIEDADES	LIQUIDEZ
	AGENTES	MÃO-DE-OBRA
	PROCESSOS	INFLAÇÃO, CORREÇÃO MONETÁRIA
Energy	OPERAÇÕES	CÂMBIO, CONTROLE
	ATIVIDADES	TRABALHO, EMPREGO
	RELAÇÕES ENTRE COISAS	TAXAS, VARIAÇÕES
Space	ESPAÇO, LUGAR	BRASIL, REGIÃO NORDESTE
Time	TEMPO	SÉCULO XX, 1989

(\*) Não foram traduzidas para o português com o objetivo de não se alterar o formato da sigla PMEST.

Entidades básicas são aquelas particularidades definidas de forma mais ampla por RANGANATHAN (PMEST) e categorias semânticas são uma conjugação das contribuições de VICKERY e de LANGRIDGE quanto à definição de facetas mais específicas. As categorias semânticas poderiam ser ainda mais divididas e subdivididas, de acordo com determinado nível de especificidade. Entretanto, a análise conceitual efetuada na investigação ficou delimitada pelas categorias semânticas descritas no esquema apresentado.

A inclusão de exemplos no esquema objetivou demonstrar como as categorias semânticas puderam auxiliar na identificação de palavras e/ou expressões que representassem conceitos pertinentes ao campo do conhecimento mais amplo abordado pelos 30 documentos integrantes da amostra, no caso, Economia. Ficou evidenciado que as categorias semânticas podem orientar e, portanto, conduzir o profissional da informação que domina determinada terminologia a uma identificação de conceitos semanticamente mais representativos nas diferentes áreas do conhecimento.

De posse do embasamento teórico indispensável, a seleção das palavras e/ou expressões que representassem conceitos pertinentes à Biblioteca do BNDES foi então efetuada.

Na análise conceitual dos documentos redigidos em inglês ou espanhol, as palavras e/ou expressões foram selecionadas sem tradução, a fim de evitar prováveis falhas na identificação de conceitos.

De acordo com os objetivos definidos para a investigação, não houve qualquer limitação quantitativa para a seleção das palavras e/ou expressões, como também não foi estabelecida qualquer relação de proporção entre a escolha da palavra e o número de páginas compreendidas pela seção do sumário, mesmo que, posteriormente, a contagem dos resultados a obter configurasse uma catalogação por assunto virtualmente exaustiva. Na análise conceitual foram selecionadas as palavras e/ou expressões julgadas relevantes para o contexto da Biblioteca do DNDES.

Palavras e/ou expressões já selecionadas mas repetidas em níveis de subordinação sobrejacentes foram desconsideradas.

Cabe ressaltar que a subjetividade foi um fator que, em certos casos, permeou a identificação e a seleção das palavras e/ou expressões durante a investigação, por ser um fator característico de qualquer processo humano de interpretação da informação e do qual não se pode subtrair. Houve momentos em que determinadas palavras e/ou expressões deram margem a diferentes interpretações conceituais; nestes casos, para dirimir dúvidas e evitar a seleção indevida de palavras e/ou expressões, prevaleceram o bom senso e a observância do contexto do Sistema de Informação examinado.

Também ocorreram pouquíssimos casos em que as palavras e/ou expressões dos títulos, subtítulos e dos sumários não foram suficientemente claras quanto à representação de conceitos, o que resultou na necessidade de se recorrer ao texto do documento



para o esclarecimento desejado, conforme previsto no item 7.1.

A seguir, são descritos a execução prática da análise conceitual e os registros da coleta de dados correspondentes.

#### 7.2.2 - Registro da análise conceitual

Em cópias xerox dos sumários dos 30 livros e monografias integrantes da amostra, foram destacadas todas as palavras, expressões e nomes próprios selecionados durante a análise conceitual. As palavras, expressões e nomes próprios destacados para cada documento foram então transcritos e anotados separadamente, por níveis de subordinação atribuídos aos títulos, subtítulos e sumários onde foram identificados. Neste momento foi efetuada a tradução para o português das palavras e expressões destacadas nos sumários redigidas em inglês ou espanhol.

A análise conceitual resultou na seleção de 795 (setecentos e noventa e cinco) palavras, expressões e nomes próprios incluídos nos títulos, subtítulos e sumários dos 30 livros e monografias integrantes da amostra. O Anexo 04 apresenta a quantificação da seleção efetuada em cada documento, por nível de subordinação. O Anexo 05 apresenta todas as palavras, expressões e nomes próprios selecionados para cada um dos 30 livros e monografias, incluindo os níveis de subordinação em que foram identificados e selecionados.

A partir das palavras, expressões e nomes próprios selecionados e anotados, foram distintamente formados os cabeçalhos de assunto destinados aos dois Sistemas de Informação abrangidos pela investigação. A constituição dos cabeçalhos de assunto obedeceu a sequência dos níveis de subordinação definidos para os títulos, subtítulos e sumários de cada documento.

Para a Biblioteca do BNDES, os cabeçalhos de assunto foram formados e anotados de acordo com as técnicas referentes à pós-coordenação, a nível de indexação e de recuperação da informação. O Anexo 06 apresenta os 772 (setecentos e setenta e dois) cabeçalhos de assunto destinados à Biblioteca do BNDES, para cada um dos 30 livros e monografias, incluindo os níveis de subordinação onde foram formados.

Para a Rede BIBLIODATA (CALCO), os cabeçalhos de assunto foram constituídos com as mesmas palavras, expressões e nomes próprios selecionados durante a análise conceitual (Anexo 05), conforme as regras estabelecidas pela própria Rede BIBLIODATA (CALCO) (10), observando-se a contextualização de conceitos, como determina a pré-coordenação, tanto a nível de indexação quanto a nível de recuperação da informação. Por esta razão, a formação dos cabeçalhos de assunto exigiu, em certos casos, a repetição de cabeçalhos tópicos ou de subcabeçalhos. O Anexo 07 apresenta os 677 (seiscentos e setenta e sete) cabeçalhos de assunto destinados à Rede BIBLIODATA (CALCO), para cada um dos 30 livros e monografias, incluindo os níveis de subordinação onde foram formados.

### 7.3 - Compatibilização dos cabeçalhos de assunto formados a partir da análise conceitual com os vocabulários controlados adotados pelos Sistemas de Informação investigados

A fim de se verificar a pertinência e a validade dos cabeçalhos de assunto formados a partir das palavras, expressões e nomes próprios selecionados durante a análise conceitual, foi necessário compatibilizá-los com os vocabulários controlados referentes à pós e à pré-coordenação, ou seja, aqueles adotados, respectivamente, pela Biblioteca do BNDES e pela Rede BIBLIODATA (CALCO).

Durante a compatibilização, foram observadas e obedecidas as regras fixadas por ambos os vocabulários controlados, principalmente aquelas referentes à sinonímia (usado para), a termos homógrafos, a notas de escopo e a termos compostos.

#### 7.3.1 - Na Biblioteca do BNDES

O vocabulário controlado adotado pela Biblioteca do BNDES (11) destina-se à observância das regras pertinentes à pós-coordenação tanto na indexação quanto na recuperação da informação. A atualização do vocabulário empregada na investigação foi datada de 23 de maio de 1990.

Os cabeçalhos de assunto formados foram individualmente compatibilizados com o vocabulário controlado. Inicialmente foram verificados os cabeçalhos coincidentes com aqueles atribuídos pela Biblioteca do BNDES e as respectivas associações. Em seguida foram anotados os cabeçalhos que poderiam ser acrescidos e as associações correspondentes, de acordo com o vocabulário controlado; é importante ressaltar que nenhum destes cabeçalhos e associações acrescidos poderiam ser idênticos aos cabeçalhos e associações coincidentes. Complementando a compatibilização, foram separados os cabeçalhos não disponíveis no vocabulário controlado (nem sob a forma de indicações remissivas), e os cabeçalhos cronológicos, visto que a Biblioteca do BNDES não utiliza a faceta da cronologia para a indexação e recuperação de informações contidas em livros e monografias.

#### 7.3.1.1 - Caracterização dos cabeçalhos formados.

Em consequência das distinções ocorridas na compatibilização acima descrita, foram constituídas as seguintes categorias de cabeçalhos de assunto a serem posteriormente analisados:

Cabeçalhos de assunto coincidentes são aqueles absolutamente idênticos aos atribuídos na indexação efetuada pela Biblioteca do BNDES.

Cabeçalhos de assunto coincidentes associados são aqueles que, conforme previsto pelo sistema automatizado, foram compulsoriamente acrescidos na indexação efetuada pela Biblioteca do BNDES.

Exemplo: No documento 04, a atribuição do cabeçalho EQUIPAMENTOS DE TELECOMUNICAÇÕES permitiu que o documento também fosse indexado sob EQUIPAMENTOS e TELECOMUNICAÇÕES.

Cabeçalhos de assunto acrescidos são os cabeçalhos que a investigação, com base no vocabulário controlado utilizado pela Biblioteca do BNDES, acrescentou como outros pontos de acesso ao conteúdo informativo dos livros e das monografias.

Cabeçalhos de assunto acrescidos associados são os cabeçalhos que, de acordo com o vocabulário controlado e com os recursos do sistema automatizado da Biblioteca do BNDES, seriam compulsoriamente outros pontos de acesso ao conteúdo informativo dos livros e das monografias.

Exemplo: No documento 20, atribuindo-se o cabeçalho AGROPECUÁRIA, o documento também seria indexado sob AGRICULTURA e PECUÁRIA.

Cabeçalhos de assunto cronológicos são aqueles identificados na categoria semântica Tempo, mas que foram distinguidos dos cabeçalhos de assunto acrescidos e coincidentes porque a Biblioteca do BNDES adota essa faceta somente para a indexação e recuperação de informações dispostas em tabelas estatísticas.

Cabeçalhos de assunto não disponíveis no vocabulário controlado são aqueles que, durante a compatibilização, não puderam ser acrescidos como outros pontos de acesso ao conteúdo informativo dos livros e das monografias, visto que não integravam o vocabulário controlado e também não possuíam termos equivalentes (usado para ou ver) para substituí-los.

Devido à diversidade de quantificações obtidas decorrentes das distinções de cabeçalhos de assunto a serem analisados, foi necessário constituir um Banco de Dados destinado ao assentamento dos cabeçalhos, objetivando, também, facilitar a apresentação dos resultados junto à Biblioteca do BNDES, assim como a elaboração dos Anexos correspondentes.

No Banco de Dados, para acompanhar a distinção e a contagem dos cabeçalhos de assunto, foi estabelecida uma codificação para cada categoria de cabeçalhos, como a seguir apresentado:

. Cabeçalhos de assunto coincidentes

Codificação: CO

. Cabeçalhos de assunto coincidentes associados

Codificação: AC

. Cabeçalhos de assunto acrescidos

Codificação: SO

. Cabeçalhos de assunto acrescidos associados

Codificação: AS

. Cabeçalhos de assunto cronológicos

Codificação: LO

Cabeçalhos de assunto não disponíveis no vocabulário controlado

Codificação: NO

Ainda no Banco de Dados, para cada cabeçalho de assunto, foram anotados o número do documento e o nível de subordinação onde foram formados.

### 7.3.2 - Na Rede BIBLIODATA (CALCO)

O vocabulário controlado adotado pela Rede BIBLIODATA (CALCO) (12) exige a observância das regras pertinentes à pré-coordenação tanto na indexação quanto na recuperação da informação. A atualização do vocabulário controlado utilizada na investigação foi a efetuada pela Biblioteca da Escola Superior de Guerra-ESG, que integra a Rede BIBLIODATA (CALCO), até maio de 1970.

Foram respeitadas as indicações normativas referentes a cada cabeçalho de assunto consultado no vocabulário, principalmente aquelas destinadas aos subcabeçalhos geográficos (DIRETO ou INDIRETO), aquelas que não permitem o uso de subcabeçalhos geográficos e aquelas que estabelecem a ordem de apresentação dos cabeçalhos tópicos e dos subcabeçalhos.

Em primeiro lugar, foram verificados os cabeçalhos coincidentes com aqueles atribuídos pela Rede BIBLIODATA (CALCO). A seguir, foram anotados os cabeçalhos que poderiam ser acrescentados de acordo com o vocabulário

controlado adotado pela Rede, e, finalizando, aqueles cabeçalhos que não estavam incorporados ao vocabulário controlado.

#### 7.3.2.1 - Caracterização dos cabeçalhos formados.

Em função das distinções ocorridas durante a compatibilização dos cabeçalhos de assunto formados a partir da análise conceitual com o vocabulário controlado adotado pela Rede BIBLIODATA (CALCO), foi necessário constituir as seguintes categorias de cabeçalhos de assunto a serem posteriormente analisados:

- a) Cabeçalhos de assunto coincidentes - aqueles absolutamente idênticos aos atribuídos pela catalogação por assunto efetuada na Rede BIBLIODATA (CALCO);
- b) Cabeçalhos de assunto acrescidos - aqueles que a investigação atribuiu, com base no vocabulário controlado utilizado pela Rede, e acrescentou como outros pontos de acesso ao conteúdo dos documentos examinados; e
- c) Cabeçalhos de assunto não disponíveis no vocabulário controlado - aqueles que não puderam ser acrescidos como outros pontos de acesso ao conteúdo informativo dos documentos por não integrarem o vocabulário controlado, mesmo sob a forma de termos equivalentes (usado para ou ver), ou então devido a restrições do vocabulário relativas à utilização de determinados subcabeçalhos.



Em cada uma das categorias acima descritas, os cabeçalhos de assunto foram destacados sob a forma de:

- . Cabeçalhos de assunto simples -- aqueles apresentados sem subcabeçalhos; e
- . Cabeçalhos de assunto compostos -- aqueles apresentados com um ou mais subcabeçalhos.

Especificamente quanto aos cabeçalhos de assunto compostos integrantes dos três grandes grupos de cabeçalhos, foi efetuado, em separado, o registro dos cabeçalhos tópicos e dos subcabeçalhos que os compuseram (os subcabeçalhos cronológicos foram distinguidos somente para os cabeçalhos de assunto acrescidos e para os coincidentes).

Esta caracterização tão detalhada dos cabeçalhos de assunto formados pela investigação para a Rede BIBLIODATA (CALCO) foi ocasionada pelo fato da Rede não ser apenas um Sistema de Informação pré-coordenado e automatizado destinado a atuar de forma cooperativa e/ou para diminuir custos operacionais dos registros das informações processadas por outros Sistemas que recuperam informações sob meios convencionais. A Rede BIBLIODATA (CALCO) é, também e principalmente, uma Base de Dados bibliográficos (13), cujos recursos de acesso por assunto prevêem procedimentos que permitem a formação de argumentos de busca usando-se não somente os cabeçalhos de assunto simples ou compostos, mas também os cabeçalhos tópicos e os subcabeçalhos. Os subcabeçalhos cronológicos são desprezados na formação dos argumentos de busca, mas podem atuar,

opcionalmente, como parâmetro de restrição (14).

Devido às distinções conferidas aos cabeçalhos de assunto a serem analisados, às quantificações obtidas e aos níveis de subordinação definidos para os títulos, subtítulos e sumários, e visando facilitar a apresentação dos resultados junto à Rede BIBLIODATA (CALCO) e a elaboração dos Anexos pertinentes, foi necessário também constituir um Banco de Dados destinado ao registro dos cabeçalhos.

No Banco de Dados, para acompanhar a distinção e a contagem dos cabeçalhos de assunto, foi estabelecida uma codificação para cada uma das categorias de cabeçalhos de assunto, como apresentado no Quadro 01, a seguir:

Quadro 01: REDE BIBLIODATA (CALCO): CODIFICAÇÕES ATRIBUÍDAS  
 AOS CABEÇALHOS DE ASSUNTO FORMADOS.

CATEGORIAS DE CABEÇALHOS DE ASSUNTO	CODIFICAÇÕES ATRIBUÍDAS		
	INCOINCIDENTES; ACRESCIDOS; NÃO CONTROLADO		
	INÃO DISPONÍVEIS; VOCABULÁRIO		
	CONTROLADO		
.Cab. assunto simples	A1	A0	D0
.Cab. assunto compostos			
.Cabeçalhos tópicos	B4	B1	E1
.Subcabeçalhos	B5	B2	E2
.Cab. comp. completos	C1	C0	F0
.Subcab. cronológicos	B6	B3	-

Também no Banco de Dados, para cada cabeçalho examinado, foram anotados o número do documento e o nível de subordinação onde foram formados.

#### 7.4 - Estudo da distribuição por páginas de texto dos níveis de subordinação dos sumários dos documentos integrantes da amostra.

Objetivando identificar prováveis relações entre o número de cabeçalhos formados através dos títulos, subtítulos e sumários e o número de páginas de texto dos documentos correspondentes aos níveis de subordinação definidos pela metodologia onde ocorresse a formação dos cabeçalhos, foi necessário examinar a proporção do número de páginas compreendido por cada nível de subordinação (no caso somente os níveis 2 a 5) em relação ao total de páginas de texto dos livros e das monografias.

Os níveis 0 e 1 foram excluídos do exame porque referiam-se, respectivamente, aos títulos e subtítulos, e à subordinação que obrigatoriamente todos os sumários apresentavam.

Nos grupos de documentos cujos sumários apresentavam níveis de subordinação de 2 até 5, foi verificado o número de ocorrências de cada um dos níveis e a contagem do número de páginas de texto correspondente ao nível examinado. Para fins de cálculo de relações proporcionais, foi previsto efetuar a contagem de páginas de texto, descontando-se do número total de páginas do documento aquelas referentes a: folhas-de-rosto, apresentações, prefácios, agradecimentos, dedicatórias, introduções (quando não subdivididas), conclusões, tabelas, anexos, listas de quadros, apêndices e referências bibliográficas. Além disso, foi também previsto calcular o

percentual que o total de páginas referentes a determinado nível de subordinação corresponderia em relação ao número total de páginas do documento.

O exame dos percentuais de páginas referentes a cada nível de subordinação seria um dado de suma importância para a posterior discussão dos resultados da investigação, uma vez que permitiria analisar, para cada nível de subordinação, relações entre o número de páginas e o número de cabeçalhos de assunto formados através dos sumários. As observações mais específicas acerca deste exame são apresentadas no item 8.4.

Os resultados referentes ao estudo e avaliação de todas as variáveis estabelecidas de acordo com a metodologia empregada na investigação são a seguir apresentados e discutidos.

CITACÕES E NOTAS

- 01 - MILSTAED, Jessica L. *Subject access systems: alternatives in design*. Orlando: Academic Press, 1984. 212 p.
- 02 - CLEVELAND, Donald B., CLEVELAND, Ana D. Indexing methods and procedures. In: \_\_\_\_\_. *Introduction to indexing and abstracting*. Littleton, Colo.: Libraries Unlimited, 1983. p.62-75.
- 03 - RAWSKI, Conrad H. The scientific study of subject literatures. *Visible Language*, v.12, p. 5-23, Winter 1977.
- 04 - MILSTAED, *op. cit.*, p. 90.
- 05 - RANGANATHAN, S.R. *Colon classification*. Bombay: Asia Publishing House, 1964. p. 1.25.
- 06 - VICKERY, B.C. *Classification and indexing in science*. 2.ed. London: Butterworths Scientific Publications, 1959. 235 p.
- 07 - \_\_\_\_\_. *La classification à facettes: guide pour la construction et l'utilisation de schémas spéciaux*. Paris: Gauthier-Villars, 1963. 60 p.
- 08 - \_\_\_\_\_. Knowledge representation: a brief review. *Journal of Documentation*, v. 42, n. 3, p. 145-159, Sep. 1986.

09 - LANGRIDGE, Derek. *Approach to classification: for students of librarianship*. London: Clive Bingley, 1973. p. 61.

10 - FUNDAÇÃO GETÚLIO VARGAS. Biblioteca Central, FUNDAÇÃO GETÚLIO VARGAS. Centro de Processamento de Dados. BIBLIODATA (CALCO). *Cabeçalhos de assunto: normas adotadas pela Rede BIBLIODATA / CALCO*. Rio de Janeiro: 1989. 1 v.

11 - Desde a sua criação, a Biblioteca do BNDES adotava a pré-coordenação como regra de representação conceitual, tanto para a indexação ou catalogação por assunto, quanto para a recuperação de informações, utilizando, para tanto, como vocabulário controlado, o "Subject Headings" da Biblioteca do Congresso dos Estados Unidos e uma relação de termos de indexação, formulada pela própria Biblioteca, derivada da indexação da coleção de estudos econômicos setoriais. Na ocasião, toda a indexação e recuperação da informação por assunto era efetuada por meios convencionais, uma vez que o estágio do sistema automatizado dos registros bibliográficos ainda não permitia a inclusão de termos de indexação. Em 1985, com o desenvolvimento do sistema de automação dos registros bibliográficos, todos os documentos a serem processados a partir de então passaram a ter, também, os termos de indexação, ou os cabeçalhos de assunto, incluídos na memória do sistema.

Deve registrar que, devido à nova Política de Indexação vigente, os técnicos da Biblioteca não foram obrigados a seguir ou a se basear nos vocabulários controlados até então utilizados pela Biblioteca, ficando a critério de

cada um indexar ou catalogar por assunto usando os vocabulários disponíveis, ou então de acordo com termos de indexação que lhes conviessem atribuir para a representação de conceitos, geralmente termos derivados dos próprios documentos. Paralelamente à inclusão dos termos de indexação na memória do sistema e por força do conseqüente descontrolo da linguagem de indexação, utilizada na Biblioteca do BNDES, o que poderia conduzir a resultados não muito favoráveis para a eficiência dos serviços prestados pela Biblioteca, foi elaborado um "Cadastro de Assuntos" dos termos de indexação até então incluídos no sistema automatizado, em ordem alfabética, sem quaisquer relações hierárquicas e/ou estruturais. A partir desse cadastro, o novo vocabulário controlado foi sendo constituído e aperfeiçoado, e já de forma criteriosa, embora ainda persistindo a liberdade permitida ao profissional da informação em atribuir o termo de indexação ou o cabeçalho de assunto que julgasse apropriado para o documento a processar, já que o "software" aceitava a inclusão de qualquer termo de indexação no "Cadastro de Assuntos". No início de 1990, devido a uma decisão técnico-administrativa, e com o desenvolvimento e o aperfeiçoamento do sistema automatizado da Biblioteca (TOSAGE), o processamento técnico dos documentos na Biblioteca do BNDES sofreu muitas alterações, dentre elas a adoção da pós-coordenação como regra de indexação e de recuperação de informações. Especificamente quanto ao "Cadastro de Assuntos", os subcabeçalhos registrados na memória do sistema foram transformados em cabeçalhos ou termos de indexação, porque foi prevista a recuperação da informação através da interseção de termos (lógica



booleana) como estratégia de busca. Um dos aperfeiçoamentos desse novo "Cadastro de Assuntos" é que ele permite, para determinados termos de indexação, a indexação "compulsória" do documento sob outros termos. Esse mecanismo é denominado de associação e seu objetivo é, não somente estender a exaustividade da indexação, mas, principalmente, auxiliar o usuário no momento da busca da informação desejada (como exemplos: perguntas mal formuladas ao sistema ou imprecisão de respostas fornecidas pelo sistema). Exemplos de associações são a seguir apresentados:

#### EQUIPAMENTOS AGROPECUÁRIOS

ASSOCIAR AGROPECUÁRIA

ASSOCIAR EQUIPAMENTOS

ASSOCIAR AGRICULTURA

ASSOCIAR CRIAÇÃO PECUÁRIA

#### PESQUISA MINERAL

ASSOCIAR PESQUISA

ASSOCIAR MINÉRIOS

ASSOCIAR MINERAIS

ASSOCIAR MINERAÇÃO

A determinação das associações para os termos incluídos no "Cadastro de Assuntos" é criteriosa, a fim de não se sobrecarregar a memória do computador, como também evitar prováveis imperfeições na precisão de recuperação de informações decorrentes da interseção de termos.

O "Cadastro de Assuntos" também dispõe de outros mecanismos indispensáveis à sua exequibilidade e aplicação como: VER TAMBÉM e VER. A Biblioteca do BNDES intenciona formular, a partir dos termos incluídos no "Cadastro de Assuntos", um *Thesaurus* que atuará não somente como um vocabulário efetivamente controlado, mas também como auxiliar na seleção e atribuição apropriadas de termos de indexação e/ou cabeçalhos de assunto, uma vez que incluirá as relações hierárquicas, estruturais e semânticas referentes a cada descritor.

12 - A Rede BIBLIODATA (CALCO) adota como vocabulário controlado uma tradução adaptada do *Subject Headings*, elaborada pela Biblioteca do Congresso dos Estados Unidos. O vocabulário controlado é acompanhado pelo Grupo de Trabalho de Cabeçalhos de Assunto, coordenado pela Biblioteca Central da FGV, do qual participam representantes de diversas Bibliotecas e Centros de Documentação cooperantes da Rede BIBLIODATA (CALCO). O Grupo elaborou e vem atualizando o "Manual CALCO/CAS", cujo objetivo básico é o estabelecimento de regras, que mantenham a coerência de critérios na formação de uma linguagem de indexação pré-coordenada, fundamentada no *Subject Headings* e respeitadas as particularidades da língua portuguesa. Diante da necessidade de se estabelecer uma linguagem de indexação adequada para as Bibliotecas e Centros de Documentação especializados, o Grupo está utilizando *Thesauri* e/ou Listas de Assunto específicas para complementar os recursos oferecidos pelo "Subject Headings", mantendo a estrutura de cabeçalhos de assunto. O vocabulário controlado possui um "software" específico,

destinado à inclusão, à manutenção e à atualização de cabeçalhos de assunto, conforme determinações do Grupo de Trabalho. O vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO) segue as regras pertinentes à pré-coordenação, tanto na indexação, quanto na recuperação das informações. Cada Biblioteca ou Centro de Informação integrante da Rede BIBLIODATA (CALCO) dispõe de um exemplar do vocabulário, apresentado sob a forma de fichas 5'x7', individuais, emitidas por computador, para cada cabeçalho, cabeçalho tópico ou subcabeçalho, visando facilitar a execução das frequentes atualizações remetidas pelo Grupo de Trabalho.

13 - FUNDAÇÃO GETÚLIO VARGAS. Centro de Processamento de Dados. Gerência de Sistemas de Acesso Público. *CALCO on-line*: manual do usuário. [Rio de Janeiro: 1989] 22 f.

14 - Provavelmente, recorrer aos cabeçalhos tópicos e aos subcabeçalhos na formação da estratégia de busca implicaria em recuperar informações com elevados índices de revocação e de imprecisão. Por natureza, tanto os cabeçalhos tópicos quanto os subcabeçalhos não oferecem, isoladamente, uma representação da contextualização de conceitos satisfatória para uma recuperação de informações precisa e eficiente. Além disso, devido à dimensão do registro de obras cadastradas na Rede BIBLIODATA (CALCO), o universo de obras a se selecionar, recorrendo-se aos cabeçalhos tópicos e aos subcabeçalhos, provocaria maiores dispêndios de tempo e de custo de processamento. Mesmo com os aspectos desfavoráveis descritos, esses componentes dos cabeçalhos de assunto compostos não deixam de se constituírem em outros pontos de

acesso ao conteúdo informativo de livros e monografias, e, portanto, não foram desconsiderados pela investigação.

## 8 - RESULTADOS

Antes de iniciar a apresentação e a discussão dos resultados da investigação, é preciso esclarecer alguns detalhes, no sentido de não se comprometer a sua compreensão.

Em primeiro lugar, e conforme previsto pelos objetivos da presente Dissertação, a demonstração e conseqüente discussão dos resultados foi orientada para dois aspectos principais:

os acréscimos decorrentes da investigação em relação à catalogação por assunto efetuada pelos Sistemas de Informação examinados; e

o potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos veiculados por livros e monografias.

Os acréscimos foram identificados e analisados a partir da quantificação dos cabeçalhos de assunto formados e admitidos pelos vocabulários controlados adotados pelos Sistemas de Informação, comparando-a percentualmente com a quantificação dos cabeçalhos atribuídos durante a catalogação por assunto executada pelos Sistemas de Informação examinados.

O potencial foi identificado e analisado a partir da quantificação de todos os cabeçalhos de assunto formados para os dois Sistemas de Informação, uma vez que eles correspondiam a todos os pontos de acesso ao conteúdo informativo dos livros e das monografias oferecidos pelos elementos estruturais selecionados para a execução da pesquisa (títulos, subtítulos e sumários).

Como previsto na Metodologia do presente trabalho e no desenrolar da pesquisa, foram constituídos 3 grupos básicos e distintos de cabeçalhos de assunto formados para os dois Sistemas de Informação examinados, a saber:

- . cabeçalhos de assunto acrescidos;
- . cabeçalhos de assunto não disponíveis no vocabulário controlado; e
- . cabeçalhos de assunto coincidentes com a catalogação por assunto efetuada pelos Sistemas de Informação examinados.

Para a demonstração dos acréscimos foram utilizadas somente as quantificações dos cabeçalhos de assunto acrescidos. Para a demonstração do potencial, foram utilizadas as quantificações dos três grupos de cabeçalhos acima mencionados.

Tanto a demonstração dos acréscimos quanto a do potencial dos títulos, subtítulos e sumários estão apresentadas e analisadas separadamente para cada um dos Grupos de documentos constituídos de acordo com as características da amostra (ver item 7.1.1), uma vez que a análise e a discussão das variáveis só deveria ser efetuada em livros e monografias com sumários

estrutural e hierarquicamente semelhantes.

Além das delimitações de acréscimos e de potencial, e dos 3 grupos de cabeçalhos de assunto estabelecidos, e devido às características e aos recursos de indexação e de recuperação de informações oferecidos pelos dois Sistemas de Informação examinados - a Biblioteca do BNDES (pós-coordenado) e a Rede BIBLIODATA (CALCO) (pré-coordenado) -, foi necessário estabelecer variáveis pouco comuns ao exame de dados quantitativos relativos a cabeçalhos de assunto destinados à recuperação de informações.

Para a Biblioteca do BNDES, além dos cabeçalhos de assunto propriamente ditos, e devido ao seu sistema automatizado possuir um mecanismo muito *sui generis* de indexação e de recuperação de informações, denominado associação, foi necessário estabelecer variáveis como os cabeçalhos de assunto associados. Houve também a necessidade de se destacar os cabeçalhos cronológicos, uma vez que a Biblioteca do BNDES não adota essa faceta para a indexação e recuperação de informações contidas em livros e monografias.

Houve necessidade de se observar, na Rede BIBLIODATA (CALCO), além dos cabeçalhos de assunto propriamente ditos (simples e compostos), os cabeçalhos tópicos e os subcabeçalhos que compuseram a formação de cabeçalhos compostos, isto por força dos recursos de recuperação de informações oferecidos pela Base de Dados bibliográficos da Rede. Os subcabeçalhos cronológicos também mereceram um destaque nas quantificações dos cabeçalhos acrescidos e coincidentes.

Na própria Metodologia definida para o presente trabalho (item 7.3), pode-se observar que as codificações atribuídas aos cabeçalhos de assunto formados pela investigação para os dois Sistemas de Informação foram exatamente as variáveis a considerar para a análise dos dados.

Os resultados serão apresentados em primeiro lugar para a pós-coordenação e, a seguir, para a pré-coordenação, sendo cada um dividido entre as quantificações relativas aos acréscimos obtidos e à demonstração do potencial dos títulos, subtítulos e sumários de livros e monografias como fontes de palavras e/ou expressões para a representação de conceitos. Cabe ressaltar que, para cada um dos recursos de indexação e de recuperação examinados (pós e pré-coordenação), os resultados foram sendo discutidos na seqüência de sua apresentação (itens 8.1 e 8.2).

Em seguida, é descrita uma comparação entre os principais resultados apurados e analisados para a pós e para a pré-coordenação (item 8.3).

Finalizando, são apresentados os resultados do estudo destinado à identificação de uma provável relação proporcional existente entre o número de páginas de texto compreendido pelos níveis de subordinação e o número de cabeçalhos de assunto formados nesses mesmos níveis de subordinação (item 8.4).



## 8.1 - Resultados relativos à pós-coordenação

De acordo com as variáveis definidas para a análise dos resultados da presente investigação junto a um Sistema de Informação pós-coordenado, em todos os cinco Grupos de documentos constituídos para os 30 livros e monografias integrantes da amostra, a pesquisa realizada junto à Biblioteca do DNDES levou à formação de 772 (setecentos e setenta e dois) cabeçalhos de assunto, assim distribuídos: 456 (quatrocentos e cinquenta e seis) cabeçalhos acrescidos e atribuídos a partir do vocabulário controlado adotado pela Biblioteca, 70 (setenta) cabeçalhos cronológicos, 54 (cinquenta e quatro) cabeçalhos coincidentes com a catalogação por assunto efetuada pela Biblioteca do DNDES e 192 (cento e noventa e dois) cabeçalhos não disponíveis no vocabulário controlado adotado. Especificamente junto aos 54 cabeçalhos de assuntos coincidentes e aos 456 cabeçalhos de assunto acrescidos, somaram-se, respectivamente, 31 (trinta e um) e 178 (cento e setenta e oito) cabeçalhos associados.

Para fins de apresentação do detalhamento quantitativo de todos os cabeçalhos acima mencionados, foram elaborados quatro Anexos (08, 09, 10 e 11), incluindo os valores apurados referentes a cada um dos 30 livros e monografias integrantes da amostra, de acordo com o Grupo de documentos a que correspondem. Esses Anexos foram estruturados de forma a incluir os valores obtidos nos níveis de subordinação examinados e os valores acumulados somados na seqüência dos níveis de subordinação.

Os valores apurados nos níveis de subordinação destinaram-se ao exame dos acréscimos ou do potencial dos títulos, subtítulos e sumários de acordo com o comportamento individual de cada nível de subordinação.

Os valores somados sequencialmente e apurados destinaram-se ao exame e à observação do comportamento dos resultados cumulativos de acordo com os níveis de subordinação.

O Anexo 08 apresenta a quantificação de todos os cabeçalhos de assunto acrescidos, incluindo aqueles atribuídos a partir do vocabulário controlado adotado pela Biblioteca do BNDES e os cabeçalhos associados correspondentes, de acordo com os Grupos de documentos.

O Anexo 09 apresenta a quantificação de todos os cabeçalhos cronológicos identificados por Grupos de documentos. A contagem destes cabeçalhos não integrou o Anexo 08 porque a Biblioteca do BNDES adota a cronologia somente para a indexação de tabelas estatísticas.

O Anexo 10 apresenta a quantificação de todos os cabeçalhos de assunto formados a partir da análise conceitual mas não disponíveis no vocabulário controlado adotado pela Biblioteca do BNDES, por Grupos de documentos.

O Anexo 11 apresenta a quantificação de todos os cabeçalhos de assunto formados mas coincidentes com a catalogação por assunto efetuada pela Biblioteca do BNDES e os cabeçalhos

coincidentes associados identificados, por Grupos de documentos.

Devido à dimensão e ao nível de detalhamento dos Anexos 08, 09, 10 e 11, durante a descrição dos resultados, os valores relativos a todas as variáveis examinadas e incluídas em Quadros referem-se ao total apurado nos níveis de subordinação para os cinco Grupos de documentos formados na amostra examinada, o mesmo ocorrendo quando foram apresentados os valores acumulados obtidos na seqüência dos níveis de subordinação.

#### 8.1.1 - Acréscimos obtidos junto à Biblioteca do BNDES

As variáveis examinadas para a determinação dos acréscimos oriundos da investigação relativos à Biblioteca do BNDES foram: os cabeçalhos de assunto atribuídos pela investigação de acordo com o vocabulário controlado adotado pela Biblioteca do BNDES e os cabeçalhos associados correspondentes. A partir dos dados quantitativos constantes no Anexo 08, essas duas variáveis foram juntamente quantificadas para cada um dos Grupos de documentos, de acordo com os níveis de subordinação pertinentes, como apresentado no Quadro 02, a seguir:

Quadro 02: BIBLIOTECA DO DNDES: QUANTIFICAÇÃO DOS  
 CADEÇALHOS DE ASSUNTO ACRESCIDOS E ASSOCIAÇÕES  
 CORRESPONDENTES, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS
1	3	34	-	-	-	-	37
2	21	95	200	-	-	-	316
3	9	30	50	13	-	-	102
4	6	20	17	8	4	-	55
5	4	4	50	35	24	7	124

De acordo com os dados do Anexo 09, verificou-se que do total de cabeçalhos e associações acrescidos (634), 71,92% corresponderam a cabeçalhos (456) e 28,08% a associações (178).

Todos os cabeçalhos de assunto acrescidos pela investigação a partir do vocabulário controlado utilizado pela Biblioteca do DNDES e as associações correspondentes (que receberam as codificações SO e AS, conforme consta no item 7.3.1.1), são apresentados no Anexo 12, para cada livro ou monografia e por nível de subordinação onde foram formados.

Os valores parciais e totais acima apresentados para cada grupo de documentos permitiram efetuar observações acerca da quantificação dos cabeçalhos formados e acrescidos.

O número de cabeçalhos acrescidos a partir do vocabulário controlado, assim como das associações correspondentes, foi bem expressivo:

- . 37 cabeçalhos e associações para o Grupo 1 de documentos;
- . 316 cabeçalhos e associações para o Grupo 2 de documentos;
- . 102 cabeçalhos e associações para o Grupo 3 de documentos;
- . 55 cabeçalhos e associações para o Grupo 4 de documentos; e
- . 124 cabeçalhos e associações para o Grupo 5 de documentos.

Partindo dos valores acima indicados, foi possível calcular a média de cabeçalhos e associações acrescidos para cada documento integrante dos Grupos estabelecidos, como apresentado no Quadro 03, a seguir:

Quadro 03: BIBLIOTECA DO BNDES: MÉDIAS DE CADEÇALHUS ACRESCIDOS E ASSOCIAÇÕES CORRESPONDENTES, POR DOCUMENTO.

GRUPOS DE DOCUMENTOS	NÚMERO DE DOCUMENTOS NO GRUPO	TOTAL DE CABEÇALHOS ACRESCIDOS E ASSOCIAÇÕES NO GRUPO	MÉDIA DE CABEÇALHOS E ASSOCIAÇÕES POR DOCUMENTO
1	4	37	9,25
2	16	316	19,75
3	6	102	17,00
4	2	55	27,50
5	2	124	62,00

Estas médias indicaram que a utilização dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos podem proporcionar uma

significativa formação de cabeçalhos de assunto, principalmente para os documentos integrantes do Grupo 5 (62 cabeçalhos e associações por documento). Também através das médias acima obtidas verificou-se que a formação de cabeçalhos e associações foi se tornando maior de acordo com a especificidade da abordagem dos temas centrais dos documentos.

Foi também possível observar a contribuição de cada nível de subordinação em relação ao total de cabeçalhos acrescidos e associados apurados em cada Grupo de documentos, como apresentado no Quadro 04, a seguir:

Quadro 04: BIBLIOTECA DO ENDES: CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS ACRESCIDOS E RESPECTIVAS ASSOCIAÇÕES, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES APURADOS NO	VALORES APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO													
		TOTAIS	NÍVEL 00		NÍVEL 01		NÍVEL 02		NÍVEL 03		NÍVEL 04		NÍVEL 05		
		OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%
1	37	3	8,11	34	91,89	-	-	-	-	-	-	-	-	-	-
2	316	21	6,65	95	30,06	200	63,29	-	-	-	-	-	-	-	-
3	102	9	8,82	30	29,41	50	49,02	13	12,75	-	-	-	-	-	-
4	55	6	10,91	20	36,37	17	30,91	8	14,54	4	7,27	-	-	-	-
5	124	4	3,23	4	3,23	50	40,32	35	28,23	24	19,35	7	5,64	-	-

De acordo com o Quadro 04, quanto aos percentuais apurados cabe ressaltar:

para o Grupo 1 de documentos: o nível 1 de subordinação, por razões óbvias, permitiu a formação de quase todos os cabeçalhos e associações (91,89%). O nível 0 (títulos e subtítulos) contribuiu com um número bem reduzido de cabeçalhos e associações (0,11%);

para o Grupo 2 de documentos: o nível 2 de subordinação foi aquele que permitiu a formação do maior número de cabeçalhos e associações (63,29%). O nível 1 forneceu um número menor de cabeçalhos e associações (30,06%) e o nível 0 (títulos e subtítulos) apenas 6,65% de cabeçalhos e associações;

para o Grupo 3 de documentos: também o nível 2 de subordinação foi aquele que permitiu a formação do maior número de cabeçalhos e associações (49,02%), restando para o nível 1, 29,41%, para o nível 3, 12,76% e para o nível 0 (títulos e subtítulos) 8,81%;

para o Grupo 4 de documentos: o nível 1 de subordinação foi aquele que permitiu a formação do maior número de cabeçalhos e associações (36,37%), o nível 2 apresentou um valor bem aproximado daquele do nível 1 (30,91%), o nível 3, 14,54%, e os níveis 1 e 0 (títulos e subtítulos) percentuais de 10,91% e de 7,27%, respectivamente; e

para o Grupo 5 de documentos: também o nível 2 de subordinação foi aquele que ofereceu a maior contribuição de cabeçalhos e associações (40,32%), o nível 3, 28,23%, o nível 4, 19,35%, o nível 5, 5,64% e os níveis 1 e 0 (títulos e subtítulos) percentuais idênticos de 3,23%.

No caso dos documentos integrantes dos Grupos 4 e 5, comparando-se com os demais Grupos, houve uma dispersão na formação dos cabeçalhos, provavelmente ocasionada pelo detalhamento dos assuntos tratados. Durante a realização da análise conceitual, foi verificado que muitos conceitos eram repetidos na seqüência dos níveis de subordinação de seus sumários.

Já para os documentos dos Grupos 2 e 3, ocorreu uma significativa concentração de formação de cabeçalhos no nível 2 (para os Grupos 2 e 3) e no nível 3 (para o Grupo 3). Para os documentos dos Grupos 4 e 5, apesar das dispersões de percentuais, as maiores concentrações de formação de cabeçalhos incidiram, respectivamente, nos níveis 1 e 2 e nos níveis 2 e 3.

Como previsto e com base nos percentuais referentes aos cabeçalhos atribuídos e associados apurados e apresentados no Quadro 04, uma outra avaliação de resultados pôde ser efetuada: a apuração dos cálculos percentuais já obtidos mas somados na seqüência dos níveis de subordinação, de acordo com os Grupos de documentos, visando apontar dados relativos à especificidade da abordagem de assuntos de acordo com a estrutura dos sumários. O Quadro 05, a seguir, apresenta todos estes somatórios.



Quadro 05: BIBLIOTECA DO BNDES: PERCENTUAIS DE CABEÇALHOS ACRESCIDOS E ASSOCIADOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO					
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5
1	8,11	100,00	-	-	-	-
2	6,65	36,71	100,00	-	-	-
3	8,82	38,23	87,25	100,00	-	-
4	10,91	47,28	78,19	92,73	100,00	-
5	3,23	6,46	46,78	75,01	94,36	100,00

Desconsiderando-se os documentos integrantes do Grupo 1, cujos sumários, por natureza, incluíam efetivamente uma única estrutura de detalhamento da abordagem do assunto tópico, foi possível constatar através dos somatórios relativos aos Grupos 2, 3, 4 e 5, que os níveis 2 e 3 de subordinação dos sumários foram aqueles que, seqüencialmente, alcançaram os maiores percentuais de formação de cabeçalhos de assunto, em especial até o nível 2 de subordinação.

Outro dado importante a considerar refere-se à contribuição dos títulos e subtítulos (nível 0 de subordinação) na formação de cabeçalhos de assunto. Ficou evidenciado que, comparativamente, os títulos e subtítulos ofereceram um número de cabeçalhos de assunto proporcionalmente bem inferior aos valores obtidos em todos os níveis de subordinação referentes aos sumários.

Após todas as observações atinentes aos cabeçalhos atribuídos e associados identificados para a Biblioteca do BNDES, coube então apresentar os efetivos resultados de acréscimos oriundos da investigação.

Como etapa preliminar da demonstração dos acréscimos, em termos percentuais, relativos à Biblioteca do BNDES, foi necessário quantificar os cabeçalhos de assunto oficialmente atribuídos por aquele Sistema de Informação para todos os documentos integrantes da amostra. Seguindo a mesma metodologia de agrupamento de livros e monografias com sumários estrutural e hierarquicamente semelhantes, o Anexo 13 contém o detalhamento quantitativo dos cabeçalhos de assunto atribuídos pela Biblioteca do BNDES e das associações permitidas pelo seu sistema automatizado para cada um dos Grupos de documentos constituídos.

A formulação do cálculo de acréscimos que os cabeçalhos e associações correspondentes formados pela investigação representariam, conjuntamente, em relação à catalogação por assunto efetuada pela Biblioteca do BNDES, compreendeu a

comparação, em termos percentuais, dos valores totais incluídos no Quadro 02 com os valores totais apurados para cada Grupo de documentos incluídos no Anexo 13.

Desta forma, os acréscimos obtidos pela investigação para a Biblioteca do BNDES são apresentados no Quadro 06, a seguir:

Quadro 06: BIBLIOTECA DO BNDES: ACRÉSCIMOS À CATALOGAÇÃO POR ASSUNTO OBTIDOS COM OS CABEÇALHOS E ASSOCIAÇÕES CORRESPONDENTES FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	CABEÇALHOS DE ASSUNTO E ASSOCIAÇÕES ATRIBUÍDOS PELA BIBL. DO BNDES	CABEÇALHOS DE ASSUNTO E ASSOCIAÇÕES ACRESCIDOS PELA INVESTIGAÇÃO	TOTAIS DE CABEÇALHOS E ASSOCIAÇÕES	ACRÉSCIMOS PERCENTUAIS
1	20	37	57	185,00 %
2	82	316	398	385,37 %
3	26	102	128	392,31 %
4	14	55	69	392,86 %
5	4	124	128	3.100,00 %

Como observações acerca dos acréscimos apontados para a Biblioteca do BNDES, cabe mencionar:

- os valores percentuais obtidos foram muito expressivos, variando de 105,00% até 3.100%, de acordo com as características dos Grupos de documentos constituídos, indicando um surpreendente acréscimo de pontos de acesso ao conteúdo informativo dos livros e monografias investigados;
- os Grupos 2, 3 e 4 apresentaram percentuais com valores muito próximos, o que indica uma possibilidade de enriquecimento uniforme de processos de indexação e de recuperação de informações para livros e monografias;
- o Grupo 1 apresentou o menor índice de acréscimos, provavelmente em função de, por natureza, só possuírem um nível de subordinação na estrutura de seus sumários, o que reduz a disponibilidade de palavras e/ou expressões representativas de conceitos; e
- o Grupo 5 foi, comparando-se com os demais, um caso realmente atípico. Reportando-se às cópias dos sumários dos documentos 1 e 2, incluídas no Anexo 02 da presente Dissertação, pode-se observar a riqueza de palavras e/ou expressões representativas de conceitos que eles contém, o que ocasionou a formação de um considerável número de cabeçalhos de assunto. Surpreende o fato de que, para esses dois documentos, a Biblioteca do BNDES tenha atribuído apenas 4 (quatro) cabeçalhos de assunto e associações para recuperar as informações por eles veiculadas, principalmente quando se verifica que o documento nº 1 é uma coletânea.

Um outro estudo acerca dos acréscimos percentuais pôde ser efetuado. Refere-se à mensuração desses acréscimos na seqüência dos níveis de subordinação definidos para cada um dos Grupos de documentos formados.

Com base nos valores incluídos no Quadro 02, foi formulada a apuração da soma acumulada dos cabeçalhos acrescidos e associações correspondentes, na seqüência dos níveis de subordinação, como apresentado no Quadro 07, a seguir:

Quadro 07: BIBLIOTECA DO BNDES: VALORES ACUMULADOS, NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, PARA OS CABEÇALHOS ACRESCIDOS E ASSOCIAÇÕES CORRESPONDENTES, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	VALORES TOTAIS APURADOS NA SEQUENCIA DOS NÍVEIS DE SUBORDINAÇÃO						TOTAIS:
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	
1	3	37	-	-	-	-	37
2	21	116	316	-	-	-	316
3	7	37	87	102	-	-	102
4	6	26	43	51	55	-	55
5	4	8	58	93	117	124	124

De posse destes valores seqüencialmente apurados, foi possível desenvolver uma avaliação referente à contribuição cumulativa de cada nível de subordinação na apuração dos acréscimos.

Em relação ao valor total dos cabeçalhos e associações atribuídos pela Biblioteca do BNDES para cada Grupo de documentos, e a partir dos valores indicados no Quadro 07, foi efetuado o cálculo percentual de acréscimos cumulativos obtidos em cada nível de subordinação, como apresentado no Quadro 08, a seguir:

Quadro 08: BIBLIOTECA DO BNDES: PERCENTUAIS DE ACRÉSCIMOS CUMULATIVOS OBTIDOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	CABEÇALHOS E ASSOCIAÇÕES	VALORES AFURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO											
		ATÉ O NÍVEL 00		ATÉ O NÍVEL 01		ATÉ O NÍVEL 02		ATÉ O NÍVEL 03		ATÉ O NÍVEL 04		ATÉ O NÍVEL 05	
		OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%
1	20	3	15,00	37	185,00	-	-	-	-	-	-	-	-
2	82	21	25,61	116	141,46	31,6	385,37	-	-	-	-	-	-
3	26	9	34,62	39	150,00	89	342,31	102	392,31	-	-	-	-
4	14	6	42,86	26	185,71	43	307,14	51	364,29	55	392,86	-	-
5	4	4	100,00	8	200,00	58	1450,00	93	2325,00	117	2925,00	124	3100,00

Os valores percentuais de acréscimos apurados para os Grupos 2, 3 e 4 foram cumulativamente homogêneos, acentuando-se a partir do nível 1 de subordinação e com uma considerável elevação no nível 2. A formação progressiva de cabeçalhos mostrou-se equilibrada para esses documentos.

Para o Grupo 1 de documentos, pela natureza estrutural de seus sumários, praticamente todos os acréscimos foram obtidos no nível 1 de subordinação.

Também para estas observações e como já mencionado, o Grupo 5 de documentos foi um caso atípico comparando-o com os demais Grupos. Verificou-se que já a partir do nível 0 de subordinação (títulos e subtítulos) os percentuais de acréscimos foram elevados. Especificamente a partir do nível 2, os percentuais cumulativos atingiram valores altíssimos, chegando-se a até 3.100,00% no nível 5 de subordinação, certamente em função do detalhamento conferido pelos autores à abordagem do tema central dos documentos, como também devido à comparação percentual dos cálculos de acréscimos ter sido realizada com o reduzido número de cabeçalhos de assunto atribuídos durante a catalogação executada pela Biblioteca do BNDES.

Ainda como instrumento da avaliação da contribuição cumulativa de cada nível de subordinação em relação aos acréscimos totais obtidos, com base nos valores indicados no Quadro 08, foi calculada a diferença de percentuais entre os níveis de subordinação, objetivando apontar os níveis que, seqüencialmente, mais contribuíram para a composição dos valores totais de acréscimos em cada Grupo de documentos. O Quadro 09, a

seguir, apresenta estes cálculos.

Quadro 09: BIBLIOTECA DO BNDES: DIFERENÇAS DE PERCENTUAIS DE ACRÉSCIMOS ENTRE OS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS	DIFERENÇA DE % DE ACRÉSCIMOS					
	ENTRE OS NÍVEIS DE SUBORDINAÇÃO					
DE	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5
DOCUMENTOS:						
1	-	170,00	-	-	-	-
2	-	115,85	243,91	-	-	-
3	-	115,38	192,31	50,00	-	-
4	-	142,85	288,43	57,15	28,57	-
5	-	100,00	1250,00	875,00	600,00	175,00

Desconsiderando-se o Grupo 1 de documentos, foi possível verificar que as diferenças de percentuais de acréscimos entre os níveis de subordinação foram expressas em valores aproximados. Até no caso do Grupo 5, com elevados percentuais de acréscimos, a diferença entre os níveis 2 e 1 foi próxima daquelas obtidas para os Grupos 2, 3 e 4, só ocorrendo um desvio bem acentuado para a diferença entre os níveis 3 e 2.



As principais constatações resultantes das duas avaliações referentes aos acréscimos expressos em valores percentuais foram:

os títulos e subtítulos, em comparação com os sumários, forneceram um número bem inferior de cabeçalhos de assunto acrescidos;

os sumários, independente dos níveis de subordinação metodologicamente definidos para as suas estruturas, permitiram alcançar um expressivo acréscimo de cabeçalhos em relação à catalogação por assunto realizada pela Biblioteca do BNDES;

foi possível constatar que os níveis de subordinação 0, 1 e 2, sequencial e cumulativamente, foram aqueles que, para todos os Grupos de documentos, proporcionaram os maiores percentuais de acréscimos. Através dos valores incluídos no Quadro 09, verifica-se que, a partir do nível 3 de subordinação, as diferenças de percentuais de acréscimos entre os níveis decaíram sensivelmente. Esta constatação pode ser corroborada através dos valores já indicados no Quadro 04 e das observações a ele correspondentes; e

os níveis 1 e 2 de subordinação foram consideravelmente ricos em palavras e/ou expressões disponíveis no vocabulário controlado adotado pela Biblioteca do BNDES. Em relação aos níveis 3, 4 e 5 de subordinação, apesar de fornecerem um grande número de palavras e/ou expressões, suas contribuições não foram tão expressivas quanto às dos demais níveis, talvez em função da especificidade do vocabulário controlado não alcançar a especificidade da abordagem dos assuntos tratados pelos documentos integrantes dos Grupos 3, 4 e 5.

Perante os resultados obtidos pela investigação realizada junto à Biblioteca do BNDES, foi possível inferir que recorrer até o nível 2 de subordinação dos sumários (não se omitindo, é certo, os títulos e subtítulos) para a identificação de palavras e/ou expressões representativas de conceitos veiculados por livros e monografias é muito proveitoso e pertinente, no sentido de se permitir que a catalogação por assunto venha a cobrir, de modo satisfatório, a especificidade que os autores conferiram ao assunto central de seus trabalhos. Os níveis 3, 4 e 5 de subordinação poderiam vir a serem utilizados em Sistemas de Informação pós-coordenados que desejassem atingir um elevado nível de precisão durante a recuperação de informações.

é evidente que a utilização de palavras e/ou expressões extraídas concomitantemente dos títulos, subtítulos e sumários ficaria a critério do Sistema de Informação em exame, e estabelecida através de uma Política de Indexação. Esta Política de Indexação, ao delimitar claramente o nível de especificidade da indexação dos documentos, poderia indicar também, caso a Biblioteca do BNDES acatasse as sugestões do presente trabalho, até de qual nível de subordinação dos sumários de livros e monografias deveriam ser extraídas as palavras e/ou expressões para utilização nos processos de indexação e de recuperação de informações.

Independente de todas as demonstrações de acréscimos relativos à Biblioteca do BNDES já descritas e analisadas, foi efetuada uma comparação entre as médias de cabeçalhos e associações atribuídos por aquele Sistema de Informação por documento e a média de cabeçalhos e associações acrescentados pela investigação também por documento (ver Quadro 3), como apresentado no Quadro 10, a seguir:

Quadro 10: BIBLIOTECA DO BNDES: COMPARAÇÃO ENTRE AS MÉDIAS DE CABEÇALHOS E ASSOCIAÇÕES ATRIBUÍDOS PELA BIBLIOTECA E AS MÉDIAS DE CABEÇALHOS E ASSOCIAÇÕES ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO.

GRUPOS	NÚMERO DE DOCUMENTOS NO GRUPO	BIBLIOTECA DO BNDES	MÉDIA DE CABEÇALHOS E ASSOCIAÇÕES ATRIBUÍDOS POR DOCUMENTO	INVESTIGAÇÃO	MÉDIA DE CABEÇALHOS E ASSOCIAÇÕES ACRESCIDOS POR DOCUMENTO	SOMA DAS MÉDIAS APURADAS
1	4	20	5,00	37	9,25	14,25
2	16	82	5,13	316	19,75	24,88
3	6	26	4,34	102	17,00	21,34
4	2	14	7,00	55	27,50	34,50
5	2	4	2,00	124	62,00	64,00

O Anexo 14 apresenta um gráfico de barras que ilustra a comparação entre as médias apuradas e incluídas no Quadro 10 acima. É importante esclarecer que as primeiras barras correspondem às médias de cabeçalhos e associações atribuídos pela Biblioteca do BNDES por documento, conforme o Grupo que integram. As barras intermediárias referem-se às médias de cabeçalhos e associações acrescidos pela investigação por documento. Comparando-se estas duas barras pode-se visualmente demonstrar os substanciais acréscimos obtidos pela pesquisa. Também observa-se através das primeiras barras que as médias de

cabeçalhos e associações atribuídos pela Biblioteca do BNDES decrescem à medida que os níveis de especificidade de abordagem de assuntos dos documentos são maiores, ocorrendo o inverso em relação às barras intermediárias. As barras posteriores correspondem à soma destas duas médias apuradas, o que, de certa forma, demonstra a capacidade de representação conceitual da linguagem de indexação adotada pela Biblioteca do BNDES, visto que alguns dos cabeçalhos e associações formados pela investigação (ver item 8.1.2.3 seguinte) foram coincidentes com cabeçalhos e associações atribuídos pela Biblioteca.

Cabe mencionar também a nível de sugestão, que a Biblioteca do BNDES deveria promover uma revisão em suas diretrizes destinadas à indexação de livros e monografias, independentemente da adoção da proposta metodológica da presente Dissertação, uma vez que os acréscimos apontados para somente 30 dos livros e monografias incluídos em seu acervo revelaram que aqueles documentos mereceriam possuir uma maior disponibilidade de acesso ao seu conteúdo informativo.

Outro fator importante a considerar: os acréscimos apurados em relação à Biblioteca do BNDES referiram-se a cabeçalhos de assunto contados individualmente. Sendo um Sistema de Informação pós-coordenado, e devido à interseção ilimitada de termos de indexação permissível pelo seu sistema automatizado, na realidade, poder-se-ia dizer que esses acréscimos foram bem maiores do que aqueles quantitativa e percentualmente já mencionados, mesmo supondo-se a ocorrência de "falsas recuperações" resultantes de uma combinação de termos não correspondentes a determinados conceitos indexados.

8.1.2 -- Potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões destinadas à indexação de livros e monografias em Sistemas de Informação pós-coordenados.

Além das quantificações de acréscimos apresentadas no item 8.1.1, para se demonstrar o potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos veiculados por livros e monografias junto a um Sistema de Informação pós-coordenado (no caso a Biblioteca do BNDES), foi necessário também computar e analisar as seguintes variáveis (cabeçalhos formados a partir da análise conceitual):

- os cabeçalhos cronológicos;
- os cabeçalhos não disponíveis no vocabulário controlado utilizado pela Biblioteca do BNDES; e
- os cabeçalhos e associações coincidentes com alguns daqueles atribuídos pela catalogação por assunto efetuada pela Biblioteca do BNDES.

Antecedendo a demonstração deste potencial, os cabeçalhos acima indicados foram individualmente quantificados e analisados.

8.1.2.1 - Os cabeçalhos cronológicos.

Pelo fato da Biblioteca do BNDES só utilizar a cronologia como faceta destinada aos processos de indexação e de recuperação de informações para tabelas estatísticas, foi preciso apresentar e quantificar isoladamente os cabeçalhos cronológicos identificados durante a análise conceitual.

A partir dos dados quantitativos incluídos no Anexo 09, os cabeçalhos cronológicos foram computados para cada um dos Grupos de documentos, de acordo com os níveis de subordinação, como mostrado no Quadro 11, a seguir:

Quadro 11: BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS CRONOLÓGICOS IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO						TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	
1	1	7	-	-	-	-	8
2	1	8	17	-	-	-	26
3	0	1	1	3	-	-	5
4	0	3	3	7	0	-	13
5	0	4	4	8	2	0	18

Todos os cabeçalhos cronológicos formados para a Biblioteca do BNDES e referentes aos 30 livros e monografias integrantes da amostra (que receberam a codificação LO, como consta no item 7.3.1.1) encontram-se discriminados no Anexo 15, para cada documento e por nível de subordinação onde foram identificados.

Através do Quadro 11, foi possível observar que os autores, ao decidirem delimitar cronologicamente a abordagem dos assuntos, recorreram aos tópicos mais específicos de seus trabalhos.

Nos títulos e subtítulos (nível 0 de subordinação) houve pouquíssimas ocorrências de cabeçalhos cronológicos, assim mesmo nos documentos cujos sumários atingiram até o nível 2 de subordinação, ou seja, com menor especificidade de abordagem que os demais.

Quanto aos outros níveis de subordinação, também através do Quadro 11, foi verificado nos documentos dos Grupos 1, 2, 3 e 4 que as incidências de cabeçalhos cronológicos foram mais acentuadas nos últimos níveis de subordinação de seus sumários. Para os documentos integrantes do Grupo 5, o nível 3 foi o que forneceu o maior número de cabeçalhos cronológicos, havendo um decréscimo a partir do nível 4 de subordinação.

Como a faceta da cronologia praticamente não é utilizada pela Biblioteca do BNDES, não foi necessário efetuar outras observações acerca dos cabeçalhos cronológicos, uma vez que apenas a sua quantificação seria suficiente para integrar a



demonstração do potencial de palavras e/ou expressões dos títulos, subtítulos e sumários.

Entretanto, em consequência das limitações de uso da cronologia vigentes na Biblioteca do BNDES, coube sugerir como um aperfeiçoamento daquele Sistema que a cronologia viesse a ser adotada para o processamento de todos os demais documentos incluídos em seu acervo, o que, por intuição, não seria tão discutível, ou mesmo inviável, visto que os recursos operacionais do sistema automatizado poderiam perfeitamente também assimilar uma indexação cronológica. A cronologia confere maiores índices de precisão à indexação e à recuperação, ainda mais se aplicada na interseção de termos de indexação, como prevê a pós-coordenação.

8.1.2.2 Os cabeçalhos não disponíveis no vocabulário controlado.

Dentre os cabeçalhos formados a partir da análise conceitual para a Biblioteca do BNDES, alguns não puderam ser incluídos na contagem dos cabeçalhos acrescidos porque não estavam incorporados ao vocabulário controlado utilizado pela Biblioteca, nem sob a forma de indicações remissivas referentes à sinonímia. Quanto aos cabeçalhos compostos por mais de uma palavra formados e incluídos nessa categoria, não foi possível desmembrá-los em cabeçalhos simples para posterior interseção, ora devido à própria inexistência de uma das palavras no vocabulário controlado ou então por conduzirem a uma indevida representação conceitual caso as palavras fossem separadas.

Com base nos dados quantitativos constantes no Anexo 10, esses cabeçalhos foram computados para cada Grupo de documentos, de acordo com os níveis de subordinação, como apresentado no Quadro 12, a seguir:

Quadro 12: BIBLIOTECA DO BNDDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO				
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
1	0	6	-	-	-
2	7	13	67	-	-
3	4	9	17	10	-
4	0	9	12	3	1
5	1	3	9	11	8

Todos os cabeçalhos de assunto formados pela investigação mas não disponíveis no vocabulário controlado adotado pela Biblioteca do BNDDES (que receberam a codificação NO, conforme consta no item 7.3.1.1) encontram-se discriminados no Anexo 16, para cada documento e por nível de subordinação onde foram formados.

Como observações acerca desta categoria de cabeçalhos, cabe destacar que, para uma avaliação da defasagem do vocabulário controlado em relação aos conceitos veiculados pelos livros e monografias integrantes da amostra, preliminarmente, foi necessário calcular a média de cabeçalhos não disponíveis por grupos de documentos, como apresentado no Quadro 13, a seguir:

Quadro 13: BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	NÚMERO DE DOCUMENTOS NO GRUPO	TOTAL DE CABEÇALHOS NÃO DISPONÍVEIS NO GRUPO	MÉDIA DE CABEÇALHOS POR DOCUMENTO
1	4	6	1,50
2	16	87	5,44
3	6	40	6,67
4	2	25	12,50
5	2	34	17,00

A partir das médias acima apontadas, foi possível demonstrar, de imediato, que quanto maior a especificidade da abordagem que os autores conferiram a seus trabalhos, menor foi a capacidade da linguagem de indexação de atender à representação de determinados conceitos. As médias apuradas variaram de 1,50 a 17,00 cabeçalhos por documento, sendo que a

partir do Grupo 4, o valor da média quase dobrou em relação ao Grupo 3.

Continuando a avaliação, foi também necessário verificar a contribuição de cada nível de subordinação em relação ao total de cabeçalhos não disponíveis apurados por Grupos de documentos, como demonstrado no Quadro 14, a seguir:

Quadro 14: BIBLIOTECA DO BNDES: CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES AFURADOS NO GRUPO	VALORES AFURADOS EM CADA NÍVEL DE SUBORDINAÇÃO											
		NÍVEL 00		NÍVEL 01		NÍVEL 02		NÍVEL 03		NÍVEL 04		NÍVEL 05	
		OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%
1	6	0	0,00	6	100,00	-	-	-	-	-	-	-	-
2	87	7	8,05	13	14,94	67	77,01	-	-	-	-	-	-
3	40	4	10,00	9	22,50	17	42,50	10	25,00	-	-	-	-
4	25	0	0,00	9	36,00	12	48,00	3	12,00	1	4,00	-	-
5	34	1	2,94	3	8,83	9	26,47	11	32,35	8	23,53	2	5,88

Conforme o Quadro 14, em relação aos percentuais apurados, cabe ressaltar:

- para os documentos do Grupo 1, todos os cabeçalhos não disponíveis foram formados no nível 1 de subordinação; no nível 0 (títulos e subtítulos) nada foi apurado;
- para os documentos integrantes dos Grupos 2, 3 e 4, a maior parte dos cabeçalhos foi obtida no nível 2 de subordinação. Especificamente nos Grupos 3 e 4, os percentuais decaíram a partir do nível 3 de subordinação. O nível 0 (títulos e subtítulos) contribuiu com valores bem inferiores e até nulos (como no Grupo 4); e
- no caso dos documentos do Grupo 5, ocorreu uma dispersão entre as contribuições dos níveis de subordinação, as quais se concentraram nos níveis 2, 3 e 4, e decaíram no nível 5.

Complementando a avaliação dos cabeçalhos não disponíveis, recorreu-se também à computação acumulada dos percentuais apurados no Quadro 14, na seqüência dos níveis de subordinação, conforme apresentado no Quadro 15, a seguir:

Quadro 15: BIBLIOTECA DO BNDES: PERCENTUAIS DE CABEÇALHOS DE ASSUNTO NÃO DISPONÍVEIS NO VOC. CONTROLADO, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO				
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
1	0,00	100,00	-	-	-
2	8,05	22,99	100,00	-	-
3	10,00	32,50	75,00	100,00	-
4	0,00	36,00	84,00	96,00	100,00
5	2,94	11,77	38,24	70,59	94,12

Através do Quadro 15, foi revelado que, cumulativamente, quase todos os cabeçalhos não disponíveis foram obtidos até o nível 2 de subordinação. Entretanto, essa constatação não pôde ser estendida aos documentos do Grupo 5, no qual, somente a partir do nível 3 de subordinação, os percentuais superaram 50% dos cabeçalhos formados.

A avaliação dos dados quantitativos referentes aos cabeçalhos não disponíveis possibilitou constatar que existe uma concentração de palavras e/ou expressões potencialmente representativas de conceitos até o nível 2 de subordinação dos sumários, considerando-se também os títulos e subtítulos (nível 0). A partir do nível 3, essa concentração é reduzida, diluindo-se conforme a especificidade da abordagem do tema

central tratado pelo trabalho. Já no ítem 8.1.1, quando foram avaliados os resultados referentes aos acréscimos apurados pela investigação, os valores acumulados até o nível 2 de subordinação também responderam pelos maiores percentuais de acréscimos.

Retomando a não disponibilidade de cabeçalhos de assunto no vocabulário controlado adotado pela Biblioteca do BNDES, ficou demonstrado que, em proporção aos cabeçalhos atribuídos acrescidos pela investigação (456), sem se considerar as associações correspondentes, a linguagem de indexação utilizada estava 42,11% aquém da desejável cobertura que deveria proporcionar em relação à terminologia do campo do conhecimento abrangido pelos documentos integrantes da amostra (Economia). Esse é um dos problemas que normalmente se verifica em Sistemas de Informação: linguagens de indexação desatualizadas, que não acompanham coerentemente o desenvolvimento da terminologia empregada na literatura sobre determinado assunto (01), e que, durante esta investigação pôde ser também detectado, até de uma maneira bem direta, uma vez que os cabeçalhos de assunto foram formados a partir de palavras utilizadas pelos autores dos livros e das monografias, ao redigirem os respectivos sumários.

0) Anexo 17 compreende a ordenação alfabética de todos os cabeçalhos de assunto não disponíveis no vocabulário controlado identificados durante a investigação (já incluídos no Anexo 16, ordenados por documento), independente do livro ou da monografia

e do nível de subordinação onde foram formados. Através deste Anexo 17 foi possível identificar cabeçalhos comuns a alguns documentos, o que justificaria a sua inclusão no vocabulário controlado utilizado pela Biblioteca do BNDES. Estes cabeçalhos são a seguir indicados:

- . ANÁLISE ECONÔMICA
- . ARRECADAÇÃO
- . CRISE FINANCEIRA
- . DESVALORIZAÇÃO CAMBIAL
- . ECONOMIA ABERTA
- . INFLAÇÃO ZERO
- . MONOPÓLIOS
- . OLIGOPÓLIOS
- . ORIGENS
- . QUESTÃO AGRÁRIA
- . REDUÇÃO

Além dos cabeçalhos não disponíveis comuns a alguns documentos, através do Anexo 17 foi possível também verificar que outros termos muito pertinentes ao campo Economia não integravam o vocabulário controlado, inclusive termos bem consolidados pela literatura. Surpreende que a "liberdade" de atribuição de cabeçalhos de assunto para posterior incorporação ao vocabulário controlado permitida pela diretriz do processo de indexação vigente na Biblioteca do BNDES não tenha, em algum momento, ocasionado a incorporação de termos tão estabelecidos para a representação de conceitos econômicos e afins.



### 8.1.2.3 - Os cabeçalhos coincidentes.

Ainda em referência aos cabeçalhos de assunto formados a partir da análise conceitual, alguns foram identificados como coincidentes com parte daqueles atribuídos durante a catalogação por assunto executada pela Biblioteca do BNDES para os 30 documentos integrantes da amostra, os quais, conseqüentemente, não integraram o grupo de cabeçalhos acrescidos pela investigação.

A partir dos dados quantitativos incluídos no Anexo 11, os cabeçalhos coincidentes e associações correspondentes foram computados para cada Grupo de documentos, de acordo com os níveis de subordinação, como consta no Quadro 16, a seguir:

Quadro 16: BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CABEÇALHOS DE ASSUNTO COINCIDENTES E ASSOCIAÇÕES CORRESPONDENTES IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO:						TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	
1	9	3	-	-	-	-	12
2	37	6	3	-	-	-	46
3	12	2	0	0	-	-	14
4	8	1	0	0	2	-	11
5	1	0	0	1	0	0	2

Todos os cabeçalhos e associações formados pela investigação mas identificados como coincidentes com a catalogação por assunto efetuada pela Biblioteca do BNDES (que receberam, respectivamente, as codificações CO e AC, conforme consta no item 7.3.1.1) encontram-se discriminados no Anexo 18, para cada livro ou monografia, e por nível de subordinação onde foram formados.

Pelo Quadro 16, foi possível observar que, em todos os Grupos de documentos, a maior parte dos cabeçalhos coincidentes foi formada no nível 0 de subordinação (títulos e subtítulos), havendo também uma relativa participação do nível 1 de subordinação. Os demais níveis praticamente não contribuíram para a formação dessa categoria de cabeçalhos.

As médias de cabeçalhos coincidentes e associações correspondentes identificados por livro ou monografia foram aproximadas para os Grupos 1, 2 e 3, mais elevada para o Grupo 4 e bem mais reduzida para o Grupo 5, como apresentado no Quadro 17, a seguir:

Quadro 17: BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS COINCIDENTES E ASSOCIAÇÕES CORRESPONDENTES, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	NÚMERO DE DOCUMENTOS NO GRUPO	TOTAL DE CABEÇALHOS COINCIDENTES E ASSOCIAÇÕES NO GRUPO	MÉDIA DE CABEÇALHOS POR DOCUMENTO
1	4	12	3,00
2	16	46	2,88
3	6	14	2,34
4	2	11	5,50
5	2	2	1,00

Comparando-se o Quadro 17 com o Quadro 03 (incluído no item B.1.1), observa-se que o Grupo 5 de documentos, onde a média de cabeçalhos e associações acrescidos por livro ou monografia foi maior - 62 -, no caso dos cabeçalhos coincidentes, a média apurada para este mesmo Grupo foi a menor de todas, revelando que os documentos com maior especificidade de abordagem de assuntos perceberam da Biblioteca do BNDES um número de cabeçalhos proporcionalmente inferior aos demais, fato que deve ter provocado a obtenção de acréscimos tão elevados no Grupo 5.

As quantificações apresentadas no Quadro 16 também podem indicar uma provável tendência dos indexadores de recorrerem aos títulos e subtítulos e à primeira hierarquização dos sumários

para determinar a representação dos conceitos veiculados pelos livros e monografias. Comparando-se o número total de cabeçalhos de assunto e de associações correspondentes conferidos pela catalogação por assunto executada pela Biblioteca do BNDES para todos os documentos integrantes da amostra - 146 cabeçalhos e associações - (ver Anexo 13), com o número total de cabeçalhos e associações identificados como coincidentes - 85 cabeçalhos e associações - , foi verificado que os cabeçalhos coincidentes correspondiam a 58,22% do total de cabeçalhos e associações atribuídos por aquele Sistema de Informação, o que, de certa forma, pode confirmar essa tendência.

#### 0.1.2.4 - O potencial dos títulos, subtítulos e sumários.

Como já mencionado, para se efetuar a demonstração do potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos em um Sistema de Informação pós-coordenado (no caso a Biblioteca do BNDES), foi preciso utilizar como variáveis a computar e a analisar conjuntamente:

- . os cabeçalhos acrescidos;
- . os cabeçalhos cronológicos;
- . os cabeçalhos não disponíveis no vocabulário controlado; e
- . os cabeçalhos coincidentes.

Especificamente para os cabeçalhos acrescidos e coincidentes, foram também computadas as associações correspondentes identificadas pela investigação.

Partindo das quantificações já apresentadas por Grupos de documentos nos itens 8.1.1, 8.1.2.1, 8.1.2.2 e 8.1.2.3, foi efetuada a soma de todos os cabeçalhos acima indicados, utilizando-se os valores incluídos nos Quadros 02, 11, 12 e 16, conforme apresentado no Quadro 18, a seguir, por Grupos de documentos e de acordo com o nível de subordinação onde foram formados.

Quadro 18: BIBLIOTECA DO BNDES: QUANTIFICAÇÃO DOS CADEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO						TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	
1	13	50	-	-	-	-	63
2	66	122	287	-	-	-	475
3	25	42	68	26	-	-	161
4	14	33	32	18	7	-	104
5	6	11	63	55	34	9	178

O total apontado de 981 cabeçalhos formados correspondeu a:

- . 634 cabeçalhos acrescidos e associações correspondentes;
- . 70 cabeçalhos cronológicos;
- . 192 cabeçalhos não disponíveis no vocabulário controlado; e
- . 85 cabeçalhos coincidentes e associações correspondentes.

Diversas análises puderam ser efetuadas em função das quantificações incluídas no Quadro 18.

As médias de pontos de acesso identificados nos títulos, subtítulos e sumários para cada documento integrante dos Grupos estabelecidos, indicaram que quanto mais hierarquizadas eram as estruturas dos sumários, maiores foram as possibilidades de formação de cabeçalhos de assunto, apesar do discreto decréscimo ocorrido do Grupo 2 para o Grupo 3, como mostrado no Quadro 19, a seguir:

Quadro 19: BIBLIOTECA DO BNDES: MÉDIAS DE CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, POR DOCUMENTO.

GRUPOS DE DOCUMENTOS	NÚMERO DE DOCUMENTOS NO GRUPO	TOTAL DE CABEÇALHOS NO GRUPO	MÉDIA DE CABEÇALHOS POR DOCUMENTO
1	4	63	15,75
2	16	475	29,69
3	6	161	26,84
4	2	104	52,00
5	2	178	89,00

Para a análise da participação de cada nível de subordinação no total de cabeçalhos formados para os Grupos de documentos, foi necessário efetuar os respectivos cálculos percentuais, como apresentado no Quadro 20, a seguir:

Quadro 20: BIBLIOTECA DO BNDES: CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS	VALORES	VALORES APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO											
		NÍVEL 00		NÍVEL 01		NÍVEL 02		NÍVEL 03		NÍVEL 04		NÍVEL 05	
DE	APURADOS	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
DOCUMENTOS	GRUPO	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %	OBTIDO: %
1	63	13   20,63	50   79,37	-	-	-	-	-	-	-	-	-	-
2	475	66   13,89	122   25,68	287   60,43	-	-	-	-	-	-	-	-	-
3	161	25   15,53	42   26,09	68   42,23	26   16,15	-	-	-	-	-	-	-	-
4	104	14   13,46	33   31,73	32   30,77	18   17,31	7   6,73	-	-	-	-	-	-	-
5	178	6   3,37	11   6,18	63   35,39	55   30,90	34   19,10	9   5,06	-	-	-	-	-	-

Com exceção do Grupo 1 de documentos, foi verificado para os demais Grupos que o nível 2 de subordinação foi aquele que permitiu a formação do maior número de cabeçalhos de assunto, apesar da discreta diferença para maior entre os níveis 1 e 2 no Grupo 4 de documentos. Observou-se também para os livros e monografias dos Grupos 3, 4 e 5, que a partir do nível 3 de subordinação, a formação de cabeçalhos decaiu. Para os documentos integrantes do Grupo 5, foi reparada uma dispersão maior na formação de cabeçalhos, concentrados, principalmente, nos níveis 2 e 3 de subordinação. Os títulos e subtítulos, comparando-os com os sumários, permitiram a formação de cabeçalhos em proporções bem inferiores.

Continuando as observações relativas à contribuição dos níveis de subordinação, somando-se os percentuais indicados no Quadro 20 na seqüência dos níveis de subordinação, foi verificado que, com exceção dos Grupos 1 e 5, recorrendo-se até o nível 2 já era permitido formar, no mínimo, 75,96% dos cabeçalhos. O Grupo 1 de documentos obteve praticamente toda a formação de cabeçalhos no nível 1, como era de se esperar. O Grupo 5 alcançou um percentual acumulado semelhante aos dos Grupos 2, 3 e 4, já no nível 3 de subordinação (75,84%). Essas observações podem ser conferidas no Quadro 21, a seguir, o qual apresenta a soma acumulada dos percentuais apurados, na seqüência dos níveis de subordinação.

Quadro 21: BIBLIOTECA DO BNDES: PERCENTUAIS DE CABEÇALHOS DE ASSUNTO FORMADOS PELA INVESTIGAÇÃO, ACUMULADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO					
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5
1	20,63	100,00	-	-	-	-
2	13,89	39,57	100,00	-	-	-
3	15,53	41,62	83,85	100,00	-	-
4	13,46	45,19	75,96	93,27	100,00	-
5	3,37	9,55	44,94	75,84	94,94	100,00



Com base nos valores apurados e nas observações já efetuadas acerca dos cabeçalhos de assunto formados pela presente Dissertação para um Sistema de Informação pós-coordenado, foi possível constatar que:

- recorrer somente aos títulos e subtítulos de livros e monografias para a identificação de palavras e/ou expressões passíveis de representação conceitual não é suficiente para permitir uma indexação consistente e também para se atingir níveis de especificidade satisfatórios e condizentes com a riqueza das informações veiculadas por esses documentos;

- houve uma expressiva formação de cabeçalhos até o nível 2 de subordinação dos sumários (acima de 70% do total), o que permitiu inferir que recorrer até esse nível de subordinação para a identificação de palavras e/ou expressões com elevado potencial de representação conceitual e também objetivando atingir maiores índices de especificidade de indexação para ampliar o acesso ao conteúdo informativo de livros e monografias, é extremamente apropriado e comprovadamente passível de adoção por parte dos Sistemas de Informação pós-coordenados, ainda mais se automatizados como é a Biblioteca do BNDES; e

os resultados obtidos junto ao Grupo 5 de documentos, o qual apontou uma dispersão maior do que os demais Grupos na formação de cabeçalhos de assunto, puderam direcionar sugestões para que, em se tratando de livros ou monografias com abordagens bem específicas de seus temas centrais, seja utilizado até o nível 3 de subordinação para a identificação das palavras.

É certo que a utilização conjunta de títulos, subtítulos e sumários para a identificação de palavras representativas de conceitos ficaria condicionada a uma Política de Indexação vigente para o Sistema de Informação, a qual, ao estipular o nível de especificidade de indexação dos documentos a processar, recomendaria recorrer-se ao nível de subordinação dos sumários que, conforme demonstrado pela presente Dissertação, fosse compatível com a referida especificidade.

Diante dos expressivos resultados referentes aos acréscimos (item 8.1.1), pode até ser questionado se a utilização de títulos, subtítulos e sumários provocaria uma indexação exaustiva e, por conseqüência, a elevação dos índices de revocação que, por sua vez, poderiam comprometer os índices de precisão e de qualidade da recuperação de informações. Estas considerações ficariam restritas ao âmbito das Políticas de Indexação que vigorariam nos Sistemas de Informação. Mais importante é que ficou demonstrado que documentos tão relevantes e tão estabelecidos na literatura especializada, como é o caso dos livros e das monografias, possuem seu conteúdo informativo muito pouco acessível para posterior recuperação.

Como já mencionado, um dos fatores que certamente ocasionam as limitações de um maior acesso ao conteúdo informativo de livros e monografias é a atribuição de um reduzido número de cabeçalhos de assunto para estes documentos em relação ao seu potencial informativo, uma rotina comumente verificada nos Sistemas de Informação, onde livros e monografias não são considerados como documentos "de ponta" como, a exemplo, os artigos de periódicos, e, portanto, não merecem um tratamento mais cuidadoso e, conseqüentemente, uma indexação mais profunda. Catálogos e listagens de Bibliotecas costumam ser eficazes para a recuperação por autores e/ou por títulos de livros e monografias. Quanto à recuperação por assunto, as deficiências das linguagens de indexação utilizadas pelos Sistemas e a pequena disponibilidade de cabeçalhos de assunto por livro ou monografia indexado concorrem para que a incidência de recuperações específicas ou exaustivas ocorra em volume inferior, como também concorrem para que os investimentos efetuados na aquisição, processamento e manutenção destes documentos não tenham o esperado retorno (02).

Em relação aos usuários dos Sistemas de Informação, as recuperações de informações veiculadas por livros e monografias bem sucedidas raramente ocorrem mediante o uso dos catálogos por assunto. Na maioria das solicitações, o usuário, se necessita de uma informação contida nestes documentos, ele já se dirige ao Sistema com o nome do(s) autor(es) ou então com o título do documento que procura. Este fato ocorre, principalmente em Bibliotecas e Centros especializados.

As informações veiculadas por livros e monografias teriam uma maior e melhor recuperação e utilização recorrendo-se com maior intensidade aos catálogos de autor ou de título do que aos catálogos por assunto? Certamente, não. Esta questão, que em geral permeia os serviços prestados por Bibliotecas e Centros de Documentação/Informação, é extremamente clara ao se observar, através dos sumários da amostra examinada (ver Anexo 2), os diferentes níveis de especificidade que os autores concederam a seus trabalhos, aos quais deveria corresponder uma indexação condizente. Também em função de suas limitações operacionais e financeiras, alguns Sistemas indexam não somente livros e monografias de modo bem superficial, prejudicando, infelizmente, a possibilidade de recuperação de informações por vezes de primordial relevância para seus usuários. Um livro ou uma monografia, em tese, deveria ter seu conteúdo muito mais explorado para posterior recuperação do que propriamente o é, haja visto a comparação entre o número de cabeçalhos de assunto que a Biblioteca do BNDES atribuiu aos 30 documentos integrantes da amostra e os percentuais de acréscimos aferidos.

Já que essa conduta necessita ser alterada no sentido de se promover o aumento de acessos mais apropriados e específicos ao conteúdo informativo de livros e monografias que, conseqüentemente, ocasionariam um maior retorno dos investimentos efetuados pelos Sistemas de Informação, a principal intenção da presente Dissertação foi justamente apresentar, como alternativa para um Sistema de Informação

pós-coordenado, um procedimento cujas principais vantagens seriam: utilização de cabeçalhos de assunto idênticos ou bem próximos da linguagem dos autores, maior facilidade para o processo de interpretação conceitual dos documentos, otimização do uso dos acervos bibliográficos, e, principalmente, a possibilidade de se fornecer aos usuários respostas mais eficazes.

## 8.2 - Resultados relativos à pré-coordenação

Conforme as variáveis estabelecidas para a análise dos resultados oriundos da presente Dissertação dirigidos a um Sistema de Informação pré-coordenado, em todos os cinco Grupos de documentos constituídos para as 30 livros e monografias integrantes da amostra, a investigação levou à formação de 677 (seiscentos e setenta e sete) cabeçalhos de assunto simples e compostos, assim especificados: 412 (quatrocentos e doze) cabeçalhos de assunto acrescidos e atribuídos com base no vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO), 236 (duzentos e trinta e seis) cabeçalhos não disponíveis no vocabulário controlado e 29 (vinte e nove) cabeçalhos coincidentes com a catalogação por assunto efetuada pela Rede.

Objetivando apresentar o detalhamento quantitativo de todos os cabeçalhos acima descritos, foram elaborados três Anexos (19, 20 e 21), os quais incluem os valores apurados correspondentes a cada um dos 30 livros e monografias integrantes da amostra, de acordo com o Grupo de documentos que integraram. Estes Anexos também foram estruturados de forma a incluir os valores obtidos nos níveis de subordinação examinados e os valores acumulados somados na seqüência dos níveis de subordinação.

Tendo sido estabelecido que para a Rede BIBLIODATA (CALCO) os resultados seriam apresentados tanto para a Rede propriamente dita quanto para a sua Base de Dados bibliográficos, de acordo com as variáveis definidas para a análise de ambos os

resultados, os Anexos 19, 20 e 21 foram constituídos de forma a incluir concomitantemente os dados quantitativos referentes a cada uma dessas variáveis, especificamente os cabeçalhos compostos, cujos cabeçalhos tópicos e subcabeçalhos foram também quantificados. Os subcabeçalhos cronológicos foram destacados dos demais subcabeçalhos devido à sua participação na recuperação de informações na Base de Dados bibliográficos ser opcional e restritiva.

Em todos estes Anexos, os valores apurados nos níveis de subordinação destinaram-se ao exame dos acréscimos ou do potencial dos títulos, subtítulos e sumários de acordo com o comportamento individual de cada nível de subordinação. Os valores somados e apurados seqüencialmente destinaram-se ao exame e à observação do comportamento dos resultados cumulativos de acordo com os níveis de subordinação.

O Anexo 19 apresenta a quantificação de todos os cabeçalhos de assunto acrescidos, atribuídos de acordo com vocabulário controlado adotado pela Rede BIBLIODATA (CALCO), por Grupos de documentos.

O Anexo 20 apresenta a quantificação de todos os cabeçalhos de assunto formados a partir da análise conceitual mas não disponíveis no vocabulário controlado utilizado pela Rede, por Grupos de documentos.

O Anexo 21 apresenta a quantificação de todos os cabeçalhos de assunto formados mas coincidentes com a catalogação por assunto efetuada pela Rede, por Grupos de documentos.

Devido à dimensão e ao nível de detalhamento dos Anexos 19, 20 e 21, durante a descrição dos resultados, os valores relativos a todas as variáveis examinadas referem-se ao total apurado para os livros e monografias integrantes dos cinco grupos formados na amostra examinada, o mesmo ocorrendo quando foram apresentados os valores acumulados obtidos na seqüência dos níveis de subordinação.

Como anteriormente mencionado, a demonstração dos resultados relativos à pré-coordenação abrangeu não somente a Rede BIBLIODATA (CALCO) mas também a sua Base de Dados bibliográficos, tanto no que se refere à apuração dos acréscimos quanto à mensuração do potencial dos títulos, subtítulos e sumários.

#### 8.2.1 - Acréscimos obtidos junto à Rede BIBLIODATA (CALCO)

As variáveis examinadas para a determinação dos acréscimos resultantes da investigação relativa à Rede BIBLIODATA (CALCO) foram os cabeçalhos de assunto simples e compostos atribuídos de acordo com o vocabulário controlado adotado pela Rede. A partir dos dados quantitativos que constam no Anexo 19, essas duas variáveis foram juntamente quantificadas para cada um dos Grupos de monografias, de acordo com os níveis de subordinação correspondentes, como apresentado no Quadro 22, a seguir:



Quadro 22 : REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS  
 CADEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS  
 ACRESCIDOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO:					TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	
1	3	23	-	-	-	26
2	15	46	124	-	-	185
3	4	11	35	12	-	62
4	2	12	9	9	6	38
5	2	5	34	35	21	101

O total de cabeçalhos acrescidos (412) correspondeu a 190 cabeçalhos simples (46,12%) e a 222 cabeçalhos compostos (53,88%).

Todos os cabeçalhos de assunto simples e compostos acrescidos pela investigação a partir do vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO) (que receberam a codificação AO e CO, conforme consta no item 7.3.2.1), são discriminados no Anexo 22, para cada documento e por nível de subordinação onde foram formados.

Baseando-se nos valores parciais e totais mostrados no Quadro 22 para cada Grupo de documentos, foi possível efetuar algumas observações relativas à quantificação dos cabeçalhos acrescidos.

O número de cabeçalhos de assunto simples e compostos acrescidos pela investigação de acordo com o vocabulário controlado foi muito significativo, como discriminado a seguir:

- . 26 cabeçalhos para o Grupo 1 de documentos;
- . 185 cabeçalhos para o Grupo 2 de documentos;
- . 62 cabeçalhos para o Grupo 3 de documentos;
- . 38 cabeçalhos para o Grupo 4 de documentos; e
- . 101 cabeçalhos para o Grupo 5 de documentos.

A partir dos valores acima indicados, foi possível calcular a média de cabeçalhos simples e compostos acrescidos para cada documento integrante dos Grupos estabelecidos, como apresentado no Quadro 23, a seguir:

Quadro 23: REDE DIBLIODATA (CALCO): MÉDIAS DE CADEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS ACRESCIDOS, POR DOCUMENTO.

GRUPOS	NÚMERO DE DOCUMENTOS	TOTAL DE CABEÇALHOS ACRESCIDOS	MÉDIA DE CABEÇALHOS POR DOCUMENTO
1	4	26	6,50
2	16	185	11,56
3	6	62	10,34
4	2	38	19,00
5	2	101	50,50

As médias calculadas indicaram uma expressiva formação de cabeçalhos de assunto disponíveis no vocabulário controlado para os livros e monografias, principalmente para aqueles integrantes do Grupo 5 (média de 50 cabeçalhos/documento), talvez em função da especificidade que os autores conferiram a seus trabalhos. Para o Grupo 4 a média foi bem inferior à registrada para o Grupo 5 (19 cabeçalhos/documento). Para os Grupos 2 e 3, as médias foram bem semelhantes, não havendo uma diferença tão acentuada quanto entre os Grupos 5 e 4. O Grupo 1 apresentou a menor média, certamente devido à pequena disponibilidade de palavras existentes em seus sumários para a formação de cabeçalhos de assunto.

Também a partir do Quadro 22, foi possível calcular a contribuição de cada nível de subordinação em relação ao total de cabeçalhos de assunto simples e compostos apurados em cada Grupo de documentos, como mostra o Quadro 24, a seguir:

Quadro 24: REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS SIMPLES E COMPOSTOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO		NÍVEL 00		NÍVEL 01		NÍVEL 02		NÍVEL 03		NÍVEL 04		NÍVEL 05	
	VALORES	TOTAIS	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%
1	26		3	11,54	23	88,46	-	-	-	-	-	-	-	-
2	185		15	8,11	46	24,86	124	67,03	-	-	-	-	-	-
3	62		4	6,45	11	17,74	35	56,45	12	19,36	-	-	-	-
4	38		2	5,27	12	31,58	9	23,68	9	23,68	6	15,79	-	-
5	101		2	1,98	5	4,96	34	33,66	35	34,65	21	20,79	4	3,96

Em função dos percentuais indicados no Quadro 24, foi possível observar:

... para o Grupo 1 de documentos: o nível 1 de subordinação foi o que mais contribuiu para a formação de cabeçalhos simples e compostos (88,46%); o nível 0 (títulos e subtítulos) permitiu a formação de um número bem inferior de cabeçalhos (11,54%); a incidência maior de formação de cabeçalhos no nível 1 pode ser justificada pelas limitações de especificidade de abordagem de assunto traduzidas pelos sumários dos livros ou das monografias integrantes do Grupo;

- para o Grupo 2 de documentos: o nível 2 de subordinação foi o que mais contribuiu para a formação de cabeçalhos (67,03% do total obtido), ocorrendo uma participação bem mais reduzida dos níveis 1 (24,86%) e 0 (8,11%);

- para o Grupo 3 de documentos: também o nível 2 de subordinação foi o que possibilitou a formação do maior número de cabeçalhos (56,45%), havendo uma queda dos percentuais no nível 3 (19,36%), no nível 1 (17,74%), e no nível 0 (6,45%);

- para o Grupo 4 de documentos: o nível 1 de subordinação contribuiu para a formação da maior parte dos cabeçalhos (31,50%), havendo uma participação idêntica dos níveis 2 e 3 (23,68%); o nível 4 forneceu cabeçalhos em proporção bem inferior (15,79%) e o nível 0 (títulos e subtítulos) apenas 5,27%;

- para o Grupo 5 de documentos: a maior parte da formação de cabeçalhos se concentrou nos níveis 2 e 3 de subordinação, com percentuais muito aproximados (33,66% e 34,65%, respectivamente); o nível 4 contribuiu com 20,79% e os demais níveis com percentuais bem inferiores, sendo que o nível 0 (títulos e subtítulos) com apenas 1,98%.

A partir dos percentuais apurados, foi possível observar que, com exceção do Grupo 1 de documentos, em todos os demais grupos, o nível 2 de subordinação foi aquele que possibilitou a formação do maior número de cabeçalhos compatíveis com o

vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO), mesmo considerando a discreta diferença de percentuais ocorrida a mais para o nível 3 no Grupo 5 de documentos.

Foi reparado que os títulos e subtítulos (nível 0 de subordinação) ofereceram pouquíssimas palavras para a formação de cabeçalhos. Verifica-se que, somente para o Grupo 1 de documentos (cujos sumários atingiam o menor nível de especificidade), o percentual de cabeçalhos formados conseguiu superar 10% do total apurado. Ainda em relação ao nível 0, foi observado que quanto mais detalhada era a estrutura dos sumários dos livros e monografias, menor foi a possibilidade de formação de cabeçalhos através dos títulos e subtítulos.

Também foi verificado que, para os documentos dos Grupos 3 e 4, a partir do nível 3 de subordinação, os percentuais foram iguais (Grupo 4) ou inferiores (Grupo 3) ao nível 2 de subordinação na formação de cabeçalhos.

No caso do Grupo 5 de documentos ocorreu uma dispersão maior do que nos outros Grupos na formação dos cabeçalhos, a qual foi mais intensa nos níveis 2 e 3 de subordinação (68,31% somados). O nível 4 forneceu sozinho 20,79% e os outros níveis (0, 1 e 5) percentuais somados iguais a 10,90% dos cabeçalhos formados. Diante desses resultados, foi possível supor que a dispersão apontada foi ocasionada pela especificidade da abordagem do tema central dos documentos, demonstrada através de seus sumários.

Ainda com base nos valores indicados no Quadro 24 e dando continuidade à análise da contribuição de cada nível de subordinação referente aos totais de cabeçalhos apurados nos grupos de documentos, cabe também apontar a soma dos percentuais obtidos na seqüência dos níveis de subordinação, como apresentado no Quadro 25, a seguir:

Quadro 25: REDE BIBLIODATA (CALCO): PERCENTUAIS DE CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS ACRÉSCIDOS, AFURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO					
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5
1	11,54	100,00	-	-	-	-
2	8,11	32,97	100,00	-	-	-
3	6,45	24,19	80,64	100,00	-	-
4	5,27	36,85	60,53	84,21	100,00	-
5	1,98	6,94	40,60	75,25	96,04	100,00

A partir dos somatórios incluídos no Quadro 25 acima, excetuando-se o Grupo 1 de documentos, cujos sumários possuíam uma única estrutura de detalhamento da abordagem do tema central dos livros e monografias, foi constatado, em todos os Grupos,

que os níveis 2 e 3 de subordinação permitiram a formação cumulativa da maior parte dos cabeçalhos disponíveis no vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO). Para os Grupos 2 e 3, até o nível 2 de subordinação houve a formação de, respectivamente, todos e quase todos os cabeçalhos. Nos Grupos 4 e 5, até o nível 3 de subordinação, foi possível formar mais de 75,00% dos cabeçalhos. Especificamente no caso do Grupo 5, devido à dispersão da formação de cabeçalhos já mencionada, somente a partir dos níveis 3 e 4 de subordinação é que a formação de cabeçalhos atingiu percentuais semelhantes aos dos Grupos 2, 3 e 4.

Foi observado também a reduzida contribuição dos títulos e subtítulos (nível 0 de subordinação) na formação de todos os cabeçalhos, comparando-a, proporcionalmente, com os percentuais apurados em todos os níveis de subordinação referentes aos sumários dos livros e monografias.

Outra observação importante diz respeito aos últimos níveis de subordinação dos documentos integrantes dos Grupos 3, 4 e 5, cujos valores individuais de cabeçalhos formados (ver e comparar os dados dos Quadros 22, 24 e 25) decaíram em relação aos níveis de subordinação anteriores.

Após todas as considerações relativas aos cabeçalhos simples e compostos formados e atribuídos com base no vocabulário controlado adotado pela Rede BIBLIODATA (CALCO), coube então apresentar os efetivos resultados de acréscimos obtidos pela presente investigação.



Como recurso necessário à demonstração dos acréscimos, em termos percentuais, referentes à Rede BIBLIODATA (CALCO), assim como à Base de Dados bibliográficos (que se encontram no item 0.2.2), foi preciso quantificar os cabeçalhos de assunto simples e compostos oficialmente atribuídos por aquele Sistema de Informação para todos os documentos integrantes da amostra; também foram quantificados, independentemente, os cabeçalhos tópicos, os subcabeçalhos e os subcabeçalhos cronológicos correspondentes. Seguindo a mesma metodologia de agrupamento de livros e monografias com sumários estrutural e hierarquicamente semelhantes, o Anexo 23 contém o detalhamento quantitativo de todos os cabeçalhos e subcabeçalhos acima mencionados atribuídos pela Rede para cada um dos documentos integrantes da amostra, de acordo com o Grupo que integraram.

A elaboração do cálculos dos acréscimos que os cabeçalhos de assunto simples e compostos formados e disponíveis no vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO) representariam, conjuntamente, em relação à catalogação por assunto executada por aquele Sistema de Informação, compreendeu a comparação, em termos percentuais, dos valores totais constantes do Quadro 22 com a soma dos valores totais dos cabeçalhos simples e compostos apurados para cada Grupo de documentos incluídos no Anexo 23.

Assim sendo, os acréscimos obtidos pela investigação para a Rede BIBLIODATA (CALCO) são apresentados no Quadro 26, a seguir:

Quadro 26 : REDE BIBLIODATA (CALCO): ACRÉSCIMOS A  
 CATALOGAÇÃO POR ASSUNTO OBTIDOS COM OS  
 CABEÇALHOS SIMPLES E COMPOSTOS FORMADOS PELA  
 INVESTIGAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS	CADEÇALHOS DE	CADEÇALHOS DE	TOTAIS	ACRÉSCIMOS
DE	ASSUNTO SIMPLES	ASSUNTO SIMPLES	DE	
	E COMPOSTOS	E COMPOSTOS	CADEÇALHOS	
	ATRIBUÍDOS	ACRESCIDOS PELA	SIMPLES	PERCENTUAIS
DOCUMENTOS	PELA REDE	INVESTIGAÇÃO	COMPOSTOS	
1	13	26	39	200,00 %
2	45	185	230	411,11 %
3	17	62	79	364,71 %
4	3	38	41	1.266,67 %
5	3	101	104	3.366,67 %

A partir dos valores percentuais indicados no Quadro 26 acima, foi possível observar que os acréscimos obtidos foram muito expressivos, oscilando entre 200,00% e 3.366,67%, conforme as características dos Grupos de documentos constituídos, denotando uma considerável ampliação do número de pontos de acesso ao conteúdo informativo dos livros e monografias.

Cada Grupo de documentos apresentou valores percentuais bem distintos.

O Grupo 1 apresentou o menor índice de acréscimos, como era de se esperar, devido à menor disponibilidade de palavras nos sumários dos documentos, mas mesmo assim os 200,00% apurados foram bem representativos. O Grupo 2 atingiu 411,11% de

acréscimos. Surpreendentemente, o Grupo 3 de documentos apresentou um percentual de acréscimos inferior ao Grupo 2 (364,71%), o que, por intuição, não deveria ter ocorrido, tendo em vista que a hierarquização dos sumários dos livros e monografias do Grupo 3 era maior do que a dos sumários dos livros e monografias do Grupo 2.

Para os Grupos 4 e 5, os percentuais de acréscimos atingiram valores surpreendentes (1.266,67% e 3.366,67%, respectivamente). Especificamente para estes Grupos, valeria a pena reportar-se às cópias dos sumários dos documentos 1, 2, 3 e 4 incluídas no Anexo 02 da presente Dissertação e observar a quantidade e a riqueza de palavras e/ou expressões que eles contém, o que certamente provocou a formação de um considerável número de cabeçalhos de assunto, em especial para as monografias 1 e 2. Foi perplexo compreender as razões que ocasionaram a atribuição de um número tão reduzido de cabeçalhos por parte da Rede BIBLIODATA (CALCO) para estes documentos, limitando sensivelmente o acesso a seus conteúdos informativos, ainda mais quando se constata que os documentos 1 e 3 são coletâneas.

Uma outra análise relativa aos acréscimos percentuais pôde ser executada: a mensuração destes acréscimos na seqüência dos níveis de subordinação estabelecidos para cada um dos Grupos de documentos formados.

Inicialmente, a partir dos valores constantes do Quadro 22, foi elaborada a apuração da soma acumulada dos cabeçalhos simples e compostos acrescidos, na seqüência dos níveis de subordinação, como apresentado no Quadro 27, a seguir:

Quadro 27: REDE BIBLIODATA (CALCO): VALORES ACUMULADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO PARA OS CABEÇALHOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO							
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS	
1	3	26	-	-	-	-	26	
2	15	61	185	-	-	-	185	
3	4	15	50	62	-	-	62	
4	2	14	23	32	38	-	38	
5	2	7	41	76	97	101	101	

Utilizando os valores indicados no Quadro 27 acima, foi possível desenvolver uma avaliação referente à contribuição cumulativa de cada nível de subordinação na apuração dos acréscimos.

Em relação ao valor total dos cabeçalhos simples e compostos atribuídos pela Rede BIBLIODATA (CALCO) para cada grupo de documentos, e a partir dos valores individuais de cada nível de subordinação incluídos no Quadro 27, foi efetuado o cálculo percentual de acréscimos cumulativos obtidos em cada nível, como apresentado no Quadro 28, a seguir:

Quadro 28: REDE BIBLIODATA (CALCO): PERCENTUAIS DE ACRÉSCIMOS CUMULATIVOS OBTIDOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	CABECALHOS DE ASSUNTO	VALORES AFURADOS NA SEQUENCIA DOS NÍVEIS DE SUBORDINAÇÃO											
		ATÉ O NÍVEL 00		ATÉ O NÍVEL 01		ATÉ O NÍVEL 02		ATÉ O NÍVEL 03		ATÉ O NÍVEL 04		ATÉ O NÍVEL 05	
DA REDE	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	
1	13	3	23,08	26	200,00	-	-	-	-	-	-	-	
2	45	15	33,34	61	135,56	185	411,11	-	-	-	-	-	
3	17	4	23,53	15	88,24	50	294,12	62	364,71	-	-	-	
4	3	2	66,67	14	466,67	23	766,67	32	1066,67	38	1266,67	-	
5	3	2	66,67	7	233,34	41	1366,67	76	2533,34	97	3233,34	101	3366,67

Os valores percentuais de acréscimos cumulativos foram muito heterogêneos para os cinco Grupos de documentos, embora, em termos gerais, tenha sido revelado o expressivo aumento de pontos de acesso ao conteúdo informativo dos livros e monografias já a partir do nível 1 de subordinação.

Os livros e monografias do Grupo 1 obtiveram a maior parte dos acréscimos no nível 1 de subordinação. A participação do nível 0 (títulos e subtítulos) foi bem inferior.

Para os documentos do Grupo 2, verificou-se que o ponto máximo de acréscimos cumulativos foi alcançado no nível 2 de subordinação. Mesmo sendo o nível 2 a última hierarquização de sumários definida para este Grupo, não se pode deixar de observar que os percentuais acumulados até o nível 1 (135,56%) não foram superiores ao valor percentual individualmente apurado no nível 2 de subordinação (275,55%).

Também para o Grupo 3, a maioria dos acréscimos percentuais cumulativos foi obtida até o nível 2 de subordinação, visto que a diferença de percentuais cumulativos entre os níveis 3 e 2 (70,59%) foi bem inferior à diferença entre os níveis 2 e 1 (205,00%).

Os livros e monografias integrantes do Grupo 4, já no nível 1 de subordinação, atingiram altíssimos índices de acréscimos percentuais cumulativos (466,67%). A participação cumulativa dos níveis 2 e 3 de subordinação foi equilibrada, considerando-se as diferenças verificadas entre os valores percentuais dos níveis 2 e 1 (300,00%) e dos níveis 3 e 2 (300,00%). A partir do nível 4 de subordinação, foi verificado um decréscimo na acumulação de percentuais, apontado pela diferença de 200,00% entre os níveis 4 e 3. Assim sendo, o ponto máximo de acréscimos ocorreu no nível 3 de subordinação. Cabe ressaltar que a obtenção de 1.266,67% de acréscimos totais foi muitíssimo expressiva.

Para o Grupo 5, até o nível 1 de subordinação, os acréscimos foram pequenos, comparando-os com o acréscimo cumulativo indicado no nível 5. Já a partir do nível 2 de

subordinação, os acréscimos cumulativos alcançaram um valor superior ao apontado no nível 4 do Grupo 4, e, na medida em que os níveis foram se sucedendo, os valores aumentaram consideravelmente. O ponto máximo de acréscimos cumulativos ocorreu no nível 3 de subordinação, o qual atingiu 2.533,34%. A partir do nível 4 de subordinação, houve uma redução na acumulação seqüencial dos acréscimos, que pode ser notada através das diferenças entre os níveis 4 e 3 (700,00%) e entre os níveis 5 e 4 (133,33%).

Um dos fatores que certamente causou a obtenção de elevados índices de acréscimos para livros e monografias integrantes do Grupo 4 e, em especial, do Grupo 5, foi o reduzidíssimo número de cabeçalhos de assunto atribuídos aos quatro documentos pela Rede BIBLIODATA (CALCO). Não se pode deixar de ressaltar que, também para os demais Grupos, os acréscimos atingiram valores muito elevados.

As principais constatações derivadas da avaliação dos acréscimos demonstrados em valores percentuais foram:

os títulos e subtítulos, comparando-os com os sumários, proporcionaram a formação de um número bem inferior de cabeçalhos de assunto simples e compostos acrescidos, o que justificou a sua pequena participação nos acréscimos cumulativos;

os sumários, independente dos níveis de subordinação metodologicamente definidos para as suas estruturas, possibilitaram o alcance de um considerável acréscimo de

cabeçalhos simples e compostos em relação à catalogação por assunto executada pela Rede BIBLIODATA (CALCO);

os níveis 1, 2 e 3 foram aqueles que, cumulativamente, permitiram atingir os maiores acréscimos percentuais cumulativos para todos os Grupos de documentos; especificamente para os Grupos 4 e 5, foi possível observar que, a partir dos níveis 4 e 5 de subordinação, o ritmo de contribuição para a formação de cabeçalhos acrescidos diminuiu; também no Quadro 22 esta redução foi observada; e

os níveis 1, 2 e 3 de subordinação, em especial os níveis 1 e 2, foram aqueles que mais contribuíram com palavras e/ou expressões para a formação de cabeçalhos de assunto; o nível 0 (títulos e subtítulos) permitiu a formação de um número bem inferior de cabeçalhos, apesar de integrar os valores percentuais cumulativos dos níveis 1, 2 e 3; quanto aos níveis 4 e 5, a redução de suas participações nos acréscimos percentuais pode ter ocorrido em função das limitações do vocabulário controlado para acompanhar a especificidade dos itens mais hierarquizados dos sumários dos documentos.

Baseando-se em todos os valores apurados e nas respectivas avaliações descritas relativas aos acréscimos obtidos pela investigação junto à Rede BIBLIODATA (CALCO), foi possível deduzir que recorrer até o nível 3 de subordinação dos sumários



para a identificação de palavras e/ou expressões representativas de conceitos veiculados em livros e monografias integrantes de acervos de Sistemas de Informação pré-coordenados é muito conveniente e adequado, no sentido de se permitir que a catalogação por assunto venha a cobrir, de modo satisfatório, a especificidade que os autores concederam a seus trabalhos. Uma outra vantagem, por sinal bem oportuna, seria a atribuição de cabeçalhos de assunto com palavras idênticas ou quase idênticas às utilizadas pelos autores, o que reduziria a probabilidade de indexações desacertadas, como também facilitaria o trabalho intelectual do profissional da informação.

A contextualização de conceitos, inerente à pré-coordenação, também deve ter provocado a obtenção de índices tão elevados de acréscimos para os documentos integrantes dos Grupos 4 e 5, isto porque, à medida que os conceitos eram identificados na seqüências dos níveis de subordinação, algumas palavras representativas anteriormente utilizadas foram obrigatoriamente repetidas para a formação de outros cabeçalhos compostos, seja sob a forma de cabeçalho tópico ou sob a forma de subcabeçalho, conforme a orientação do vocabulário controlado relativa ao cabeçalho formado. Recorrer aos níveis 4 e 5 de subordinação dos sumários para a identificação de palavras conceitualmente representativas ficaria a critério dos Sistemas de Informação interessados.

A propósito, a utilização cumulativa de qualquer um dos níveis de subordinação estipulados para os títulos, subtítulos e sumários defendida pela presente Dissertação seria vinculada a uma Política de Indexação estabelecida para Sistemas de Informação com processos de indexação e de recuperação de informações pré-coordenados, a qual definiria os graus de exaustividade ou de especificidade da indexação de livros e monografias, de acordo com seus objetivos, usuários e qualidade do acervo bibliográfico.

Como um desfecho para a demonstração dos acréscimos obtidos para a Rede BIBLIODATA (CALCO), foi efetuada uma comparação entre a média de cabeçalhos simples e compostos atribuídos por aquele Sistema de Informação por documento e a média de cabeçalhos simples e compostos acrescidos pela investigação também por documento (ver Quadro 23), como apresentado no Quadro 29, a seguir:

Quadro 29 -- REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE AS MÉDIAS DE CABEÇALHOS ATRIBUÍDOS PELA REDE E AS MÉDIAS DE CABEÇALHOS ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO.

GRUPOS	NÚMERO DE DOCUMENTOS NO GRUPO	REDE BIBLIODATA (CALCO) ATRIBUÍDOS POR DOCUMENTO	MÉDIA DE CABEÇALHOS	INVESTIGAÇÃO ACRESCIDOS POR DOCUMENTO	MÉDIA DE CABEÇALHOS	SOMA DAS MÉDIAS APURADAS
1	4	13	3,25	26	6,50	9,75
2	16	45	2,81	185	11,56	14,37
3	6	17	2,84	62	10,34	13,18
4	2	3	1,50	38	19,00	20,50
5	2	3	1,50	101	50,50	52,00

(1) Anexo 24 apresenta um gráfico de barras que ilustra a comparação entre as médias apuradas e incluídas no Quadro 29, acima. É importante esclarecer que as primeiras barras correspondem às médias de cabeçalhos de assunto atribuídos pela Rede BIBLIODATA (CALCO) por documento, conforme o Grupo que integram. As barras intermediárias referem-se às médias de cabeçalhos de assunto acrescidos pela investigação por documento. Somente comparando-se estas duas barras pode-se visualmente demonstrar os consideráveis acréscimos obtidos pela pesquisa. Também verifica-se através das primeiras barras que as médias de cabeçalhos de assunto atribuídos pela Rede BIBLIODATA (CALCO) decrescem à medida em que os níveis de especificidade de

abordagem de assuntos dos documentos são maiores, ocorrendo o inverso em relação às barras intermediárias. As barras posteriores correspondem à soma destas duas médias apuradas, o que, de certa forma, demonstra a capacidade de representação conceitual do vocabulário controlado adotado pela Rede, uma vez que alguns dos cabeçalhos de assunto formados pela investigação (ver item 8.2.3.1.2 seguinte) foram coincidentes com cabeçalhos atribuídos pela Rede BIBLIODATA (CALCO).

#### 8.2.2 - Acréscimos obtidos junto à Base de Dados bibliográficos da Rede BIBLIODATA (CALCO).

As variáveis examinadas para a determinação dos acréscimos oriundos da presente investigação junto à Base de Dados da Rede BIBLIODATA (CALCO) foram: os cabeçalhos simples e compostos já quantificados e analisados no item 8.2.1, e os cabeçalhos tópicos, subcabeçalhos e subcabeçalhos cronológicos que compuseram a formação dos cabeçalhos compostos acrescidos.

Como os subcabeçalhos cronológicos atuariam somente como parâmetro de restrição de busca na Base de Dados, foi necessário apresentar os resultados dos acréscimos excluindo-se e incluindo-se estes subcabeçalhos.

Todas as quantificações referentes aos totais apurados em cada Grupo de documentos incluídas nos itens 8.2.2.1 e 8.2.2.2 subsequentes encontram-se detalhadas, para todos os livros e monografias integrantes da amostra, no Anexo 19.

8.2.2.1 - Acréscimos sem os subcabeçalhos cronológicos.

Os cabeçalhos tópicos e os subcabeçalhos que integraram a formação dos cabeçalhos compostos acrescidos pela investigação foram conjuntamente quantificados no total, por Grupos de documentos, de acordo com os níveis de subordinação, como apresentado no Quadro 30, a seguir:

Quadro 30 : REDE - BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CADEÇALHOS TÓPICOS E SUBCADEÇALHOS INTEGRANTES DOS CADEÇALHOS COMPOSTOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO:							
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS:	
1	3	9	-	-	-	-	12	
2	11	32	70	-	-	-	113	
3	4	6	17	11	-	-	38	
4	0	11	5	3	1	-	20	
5	2	2	22	13	4	1	44	

Todos os cabeçalhos tópicos e subcabeçalhos que compuseram os cabeçalhos compostos acrescidos (que receberam, respectivamente, as codificações B1 e B2, como consta no item 7.3.2.1) encontram-se discriminados no Anexo 25, por documento e incluindo o nível de subordinação onde foram identificados.

Complementando a quantificação das variáveis destinadas ao exame dos acréscimos junto à Base de Dados da Rede BIBLIODATA (CALCO), os valores apurados no Quadro 30 acima foram somados aos valores constantes no Quadro 22 (cabeçalhos simples e compostos acrescidos), como apresentado no Quadro 31, a seguir:

Quadro 31 : REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS SIMPLES E COMPOSTOS ACRESCIDOS E DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO DE					TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	
1	6	32	-	-	-	38
2	26	78	194	-	-	298
3	8	17	52	23	-	100
4	2	23	14	12	7	58
5	4	7	56	48	25	140

Para a demonstração dos acréscimos obtidos pela investigação em relação à Base de Dados da Rede, sem se considerar os subcabeçalhos cronológicos, foi necessário computar, através do Anexo 23, os valores totais por Grupos de documentos das mesmas variáveis acima discriminadas e efetuar os cálculos percentuais sobre os valores totais indicados no Quadro 31. Os resultados destes cálculos encontram-se no Quadro 32, a seguir:

Quadro 32: REDE BIBLIODATA (CALCO): ACRÉSCIMOS À BASE DE DADOS BIBLIOGRÁFICOS, SEM OS SUBCABEÇALHOS CRONOLÓGICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	CAB. SIMPLES E COMPOSTOS, DE SUBCABEÇALHOS ATRIBUÍDOS PELA REDE	CAB. SIMPLES E COMPOSTOS, DE SUBCABEÇALHOS ACRESCIDOS PELA INVESTIGAÇÃO	TOTAIS DE CAB. SIMPLES E COMPOSTOS, DE SUBCABEÇALHOS	ACRÉSCIMOS PERCENTUAIS
1	31	38	69	122,58 %
2	102	298	400	292,16 %
3	36	100	136	277,78 %
4	8	58	66	725,00 %
5	9	145	154	1.611,12 %

Comparando-se os acréscimos percentuais obtidos pela investigação junto à Rede BIBLIODATA (CALCO) (Quadro 26) com os acréscimos percentuais apurados para a sua Base de Dados bibliográficos (Quadro 32), foi verificado que estes acréscimos foram inferiores, como demonstrado no Quadro 33, a seguir.

Quadro 33: REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE OS ACRÉSCIMOS OBTIDOS PARA A REDE E PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	ACRÉSCIMOS NA REDE (EM %)	ACRÉSCIMOS NA BASE DE DADOS (EM %)	DIFERENÇAS (EM %)
1	200,00	122,58	77,42
2	411,11	292,16	118,95
3	364,71	277,78	86,93
4	1.266,67	725,00	541,67
5	3.366,67	1.611,12	1.755,55

Pela comparação acima apresentada, foi constatado que os acréscimos à Base de Dados bibliográficos não foram tão expressivos quanto os acréscimos referentes à Rede BIBLIODATA (CALCO). Uma explicação plausível seria a freqüente repetição de palavras ou de expressões na formação dos cabeçalhos de assunto, ora como subcabeçalhos ora como cabeçalhos de assunto simples. Esta repetição de palavras conferiu uma maior disponibilidade de pontos de acesso ao conteúdo informativo dos livros e



monografias, tanto para a quantificação de acréscimos apurada pela investigação quanto para a quantificação obtida na catalogação por assunto efetuada pela Rede. Em paralelo, observou-se através do Quadro 32 que os cálculos de percentuais foram executados com valores relativamente mais aproximados do que aqueles incluídos no Quadro 26, fato que, por lógica, deve ter ocasionado a obtenção de valores percentuais inferiores.

Mesmo assim, não se pode deixar de mencionar que os acréscimos indicados para a Base de Dados bibliográficos foram consideráveis, possibilitando a ampliação dos pontos de acesso ao conteúdo informativo dos livros e monografias.

#### 0.2.2.2 - Acréscimos com os subcabecalhos cronológicos.

Continuando a avaliação dos acréscimos referentes à Base de Dados bibliográficos, foi necessário efetuar, em separado, os cálculos de acréscimos com a inclusão dos subcabecalhos cronológicos, visto que, como já mencionado, estes subcabecalhos atuariam somente como parâmetro de restrição de busca na Base.

Conseqüentemente, os acréscimos a seguir demonstrados corresponderam a todos os pontos de acesso ao conteúdo informativo dos documentos examinados permissíveis pela Base de Dados da Rede BIBLIODATA (CALCO).

Os subcabecalhos cronológicos acrescentados pela investigação foram quantificados por Grupos de documentos, como apresentado no Quadro 34, a seguir:

Quadro 34: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS SUBCABEÇALHOS CRONOLÓGICOS ACRESCIDOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO:						TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	
1	1	7	-	-	-	-	8
2	1	7	17	-	-	-	25
3	0	1	1	3	-	-	5
4	0	3	3	7	0	-	13
5	0	3	5	7	2	0	17

Todos os subcabecalhos cronológicos acrescentados (que recebem a codificação B3, como consta no item 7.3.2.1) são discriminados no Anexo 26, por documento e incluindo o nível de subordinação onde foram formados.

Através do Quadro 34, foi possível observar que os autores conferiram a abordagem cronológica dos assuntos versados em seus trabalhos nos itens mais específicos: raros foram os subcabecalhos cronológicos acrescentados através dos títulos e subtítulos.

Para a elaboração do cálculo dos acréscimos percentuais, foi necessário somar os valores incluídos no Quadro 31 e no Quadro 34, como apresentado no Quadro 35, a seguir:

Quadro 35: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DE TODOS OS CABEÇALHOS E SUBCABEÇALHOS ACRESCIDOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO DE					TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	
1	7	39	-	-	-	46
2	27	85	211	-	-	323
3	8	18	53	26	-	105
4	2	26	17	19	7	71
5	4	10	61	55	27	162

A exemplo do que ocorreu nos resultados relativos à Rede BIBLIODATA (CALCO), as médias de pontos de acesso acrescentados à sua Base de Dados apuradas por livro ou monografia indicaram que os acréscimos foram tornando-se maiores de acordo com a especificidade da abordagem dos temas centrais dos documentos, como demonstrado no Quadro 36, a seguir:

Quadro 36: REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS E SUBCABEÇALHOS ACRESCIDOS À BASE DE DADOS, POR DOCUMENTO.

GRUPOS	NÚMERO DE	TOTAL	MÉDIA
DE	DOCUMENTOS	DE PONTOS	DE PONTOS
DOCUMENTOS	NO GRUPO	NO GRUPO	DOCUMENTO
1	4	46	11,50
2	16	323	20,19
3	6	105	17,50
4	2	71	35,50
5	2	162	81,00

O cálculo dos acréscimos percentuais da Base de Dados bibliográficos incluiu os valores totais apurados no Quadro 35 e os valores totais relativos a todos os cabeçalhos e subcabeçalhos quantificados no Anexo 23, por Grupos de documentos, como apresentado no Quadro 37, a seguir:

Quadro 37: REDE BIBLIODATA (CALCO): ACRÉSCIMOS TOTAIS À BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS	PONTOS DE ACESSO	PONTOS DE ACESSO ACRESCIDOS PELA REDE	TOTAIS DE ACESSO	ACRÉSCIMOS DE PERCENTUAIS
1	36	46	82	127,78 %
2	106	323	429	304,72 %
3	36	105	141	291,67 %
4	8	71	79	887,50 %
5	10	162	172	1.620,00 %

Também para a Base de Dados bibliográficos foi importante observar e calcular a contribuição seqüencial e cumulativa dos níveis de subordinação na obtenção dos acréscimos totais de cada Grupo de documentos. A partir dos valores discriminados no Quadro 35, foi efetuada a soma dos cabeçalhos e subcabeçalhos acrescidos, na seqüência dos níveis de subordinação, como apresentado no Quadro 38, a seguir:

Quadro 38: REDE BIBLIODATA (CALCO): VALORES ACUMULADOS, NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, PARA OS CABEÇALHOS E SUBCABEÇALHOS ACRESCIDOS À BASE DE DADOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO				
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
1	7	46	-	-	-
2	27	112	323	-	-
3	8	26	79	105	-
4	2	28	45	64	71
5	4	14	75	130	157

A partir dos valores acima apurados foi efetuado o cálculo dos acréscimos percentuais cumulativos obtidos na seqüência dos níveis de subordinação, os quais são apresentados no Quadro 39, a seguir:

Quadro 39: REDE BIBLIODATA (CALCO): PERCENTUAIS DE ACRÉSCIMOS CUMULATIVOS OBTIDOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO PARA A BASE DE DADOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	TOTAIS DE PONTOS DE ACESSO FORMADOS	VALORES APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO											
		ATÉ O NÍVEL 0		ATÉ O NÍVEL 1		ATÉ O NÍVEL 2		ATÉ O NÍVEL 3		ATÉ O NÍVEL 4		ATÉ O NÍVEL 5	
DA REDE		OBTIDO	% AUM.	OBTIDO	% AUM.	OBTIDO	% AUM.	OBTIDO	% AUM.	OBTIDO	% AUM.	OBTIDO	% AUM.
1	36	7	19,44	46	127,78	-	-	-	-	-	-	-	-
2	106	27	25,47	112	105,66	323	304,72	-	-	-	-	-	-
3	36	8	22,23	26	72,23	79	219,44	105	291,67	-	-	-	-
4	8	2	25,00	28	350,00	45	562,50	64	800,00	71	887,50	-	-
5	10	4	40,00	14	140,00	75	750,00	130	1300,00	157	1570,00	162	1620,00

De acordo com os percentuais apontados pelo Quadro 39, verificou-se que a maior parte dos acréscimos foi obtida até o nível 2 de subordinação (para os livros e monografias dos Grupos 1, 2 e 3) e até o nível 3 de subordinação (para os livros e monografias dos Grupos 4 e 5). Esta observação foi absolutamente idêntica à efetuada para os valores relativos aos acréscimos percentuais cumulativos dirigidos à Rede BIBLIODATA (CALCO). Também idêntica foi a constatação de que os títulos e subtítulos pouco contribuíram na promoção dos acréscimos, ficando os sumários com o maior percentual.

A avaliação dos resultados referentes à Base de Dados bibliográficos veio corroborar todas as constatações e deduções já descritas durante a apresentação dos resultados atinentes aos acréscimos obtidos junto à Rede BIBLIODATA (CALCO), principalmente a de que recorrer até os níveis 2 e 3 de subordinação de sumários de livros e monografias para a identificação de palavras representativas de conceitos pertinentes a Sistemas de Informação pré-coordenados é consideravelmente válido e proveitoso, no sentido de se maximizar tanto a disponibilidade de pontos de acesso ao conteúdo informativo dos documentos quanto a sua utilização.

Como um complemento às observações já descritas, foi efetuada uma comparação entre as médias de pontos de acesso por documento observadas pela investigação para a Base de Dados bibliográficos, a saber: as médias relativas à catalogação executada pela Rede BIBLIODATA (CALCO) e as médias referentes aos acréscimos resultantes da investigação. As diferenças entre as médias de pontos de acesso por documento apuradas foram muito acentuadas, principalmente para aqueles integrantes dos Grupos 4 e 5, este em especial. Foi observado que o Grupo de documentos com menor especificidade de abordagem do tema central (o Grupo 1), na catalogação efetuada pela Rede recebeu um número de pontos de acesso proporcionalmente maior que os demais Grupos, cujas abordagens eram sensivelmente mais detalhadas, como demonstrado pelo Quadro 40, a seguir:



Quadro 40: REDE BIBLIODATA (CALCO): COMPARAÇÃO ENTRE AS MÉDIAS DE PONTOS DE ACESSO À BASE DE DADOS ATRIBUÍDOS PELA REDE E AS MÉDIAS DE PONTOS DE ACESSO ACRESCIDOS PELA INVESTIGAÇÃO, POR DOCUMENTO.

GRUPOS	NÚMERO DE DOCUMENTOS	BASE DE DADOS	MÉDIA DE PONTOS DE ACESSO DE DOCUMENTOS DISPONÍVEIS NA BASE	INVESTIGAÇÃO	MÉDIA DE PONTOS DE ACESSO ACRESCIDOS POR DOCUMENTO	SOMA DAS MÉDIAS
1	4	36	9,00	46	11,50	20,50
2	16	106	6,63	323	20,19	26,82
3	6	36	6,00	105	17,50	23,50
4	2	8	4,00	71	35,50	39,50
5	2	10	5,00	162	81,00	86,00

O gráfico de barras incluído no Anexo 27 ilustra a comparação entre estas médias. É importante ressaltar que as primeiras barras correspondem às médias de pontos de acesso já disponíveis na Base de Dados, por documento, em função da catalogação por assunto executada pela Rede BIBLIODATA (CALCO). As barras intermediárias referem-se às médias de pontos de acesso acrescidos pela investigação por documento. Comparando-se estas duas barras pode-se visualmente demonstrar os substanciais acréscimos obtidos pela pesquisa, principalmente a partir do Grupo 2 de documentos. Também verifica-se através das primeiras barras que as médias de pontos de acesso disponíveis na Base de

Dados decrescem à medida em que os níveis de especificidade de abordagem de assuntos dos documentos são maiores, ocorrendo o inverso em relação às barras intermediárias. As barras posteriores correspondem à soma destas duas médias apuradas, o que, também e de certa forma, demonstra a capacidade de representação conceitual do vocabulário controlado adotado pela Rede, uma vez que alguns dos cabeçalhos de assunto formados pela investigação e, por consequência, os cabeçalhos tópicos e subcabeçalhos que os integraram, foram coincidentes com cabeçalhos de assunto atribuídos pela Rede BIBLIODATA (CALCO).

### 8.2.3 - Potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões destinadas à indexação de livros e monografias em Sistemas de Informação pré-coordenados.

Para se demonstrar o potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos veiculados por livros e monografias junto a um Sistema de Informação pré-coordenado (no caso a Rede BIBLIODATA (CALCO) e a sua Base de Dados bibliográficos), foi preciso computar e analisar as seguintes variáveis (cabeçalhos formados a partir da análise conceitual):

#### a) Para a Rede BIBLIODATA (CALCO):

Os cabeçalhos simples e compostos acrescidos (já quantificados e analisados através do Quadro 22),

Cabeçalhos simples e compostos não disponíveis no vocabulário controlado adotado pela Rede, e

Cabeçalhos simples e compostos coincidentes com alguns daqueles atribuídos pela catalogação por assunto executada pela Rede.

b) Para a Base de Dados bibliográficos:

Todas as variáveis discriminadas na alínea "a" acima, e ainda,

Cabeçalhos tópicos e subcabeçalhos que compuseram a formação dos cabeçalhos compostos não disponíveis no vocabulário controlado adotado pela Rede,

Os subcabeçalhos cronológicos acrescidos (já quantificados e analisados através do Quadro 34), e

Cabeçalhos tópicos, subcabeçalhos e subcabeçalhos cronológicos que integraram a formação dos cabeçalhos compostos coincidentes com a catalogação por assunto efetuada pela Rede.

#### 8.2.3.1 - O potencial na Rede BIBLIODATA (CALCO).

Antecedendo a demonstração desse potencial, as variáveis discriminadas na alínea "a" do item 8.2.3 foram individualmente quantificadas e analisadas, com exceção dos cabeçalhos simples e compostos acrescidos (ver item 8.2.1).

8.2.3.1.1 - Os cabeçalhos simples e compostos não disponíveis no vocabulário controlado.

De acordo com os dados quantitativos incluídos no Anexo 20, os cabeçalhos simples e compostos formados a partir da análise conceitual mas não disponíveis no vocabulário controlado adotado pela Rede BIBLIODATA (CALCO), nem através de indicações remissivas, foram computados para cada um dos Grupos de documentos, de acordo com os níveis de subordinação, como apresentado no Quadro 41, a seguir:

Quadro 41: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CADEÇALHOS SIMPLES E COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO:							
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS	
1	0	11	-	-	-	-	11	
2	10	20	75	-	-	-	105	
3	5	12	17	11	-	-	45	
4	1	12	14	9	2	-	37	
5	1	4	8	15	8	2	38	

Do total de cabeçalhos de assunto acima quantificados (236 cabeçalhos), 41,53% corresponderam a cabeçalhos simples e 58,47% a cabeçalhos compostos.

Todos os cabeçalhos simples e compostos formados a partir da análise conceitual mas não disponíveis no vocabulário controlado adotado pela Rede BIBLIODATA (CALCO) (que receberam, respectivamente, as codificações DO e FO, como consta no item 7.3.2.1) encontram-se discriminados no Anexo 28, por documento, e incluindo os níveis de subordinação onde foram formados.

Através do Quadro 41, foi possível observar que a incidência da maior parte desta categoria de cabeçalhos ocorreu nos níveis de subordinação correspondentes às abordagens mais específicas dos livros e monografias, o que configurou uma provável deficiência do vocabulário controlado utilizado pela Rede BIBLIODATA (CALCO) no que se refere ao acompanhamento e à cobertura da terminologia empregada na literatura do campo Economia.

Esta deficiência pode ser ainda mais demonstrada através das médias de cabeçalhos simples e compostos não disponíveis no vocabulário controlado apuradas por documento, como apresentado no Quadro 42, a seguir:

Quadro 42: REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS SIMPLES E COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR DOCUMENTO.

GRUPOS	NÚMERO DE DOCUMENTOS	TOTAL DE CABEÇALHOS NÃO DISPONÍVEIS NO GRUPO	MÉDIA DE CABEÇALHOS POR DOCUMENTO
1	4	11	2,75
2	16	105	6,56
3	6	45	7,50
4	2	37	18,50
5	2	38	19,00

De acordo com o Quadro 42, foi verificado, pelas médias apuradas, que o número de cabeçalhos simples e compostos não disponíveis no vocabulário controlado foi aumentando conforme a especificidade da abordagem dos temas centrais dos livros e monografias. Ocorreu uma variação entre 2,75 cabeçalhos (Grupo 1) a até 19,00 cabeçalhos por documento (Grupo 5). Este indicador foi bem expressivo, no sentido de que, sendo o vocabulário controlado adotado pela Rede BIBLIODATA (CALCO) uma versão e uma adaptação do LCSH (Library of Congress Subject Headings) -- uma linguagem de indexação que intenciona abarcar, representativamente, todos os campos do conhecimento --, a desejável cobertura de seus termos, pelo menos para o campo Economia e para a amostra de documentos investigada, não foi alcançada.

Uma outra avaliação pôde ser efetuada quanto à defasagem do vocabulário controlado. Refere-se à quantificação cumulativa dos cabeçalhos de assunto simples e compostos não disponíveis no vocabulário controlado, na seqüência dos níveis de subordinação, e dos respectivos cálculos percentuais, como apresentado no Quadro 43, a seguir:

Quadro 43: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO E PERCENTUAIS DE CABEÇALHOS SIMPLES E COMPOSTOS NÃO DISPONÍVEIS NO VOC. CONTROLADO, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS	TOTAIS DE CAB. ASSUNTO	VALORES APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO												
		ATÉ O NÍVEL 0		ATÉ O NÍVEL 1		ATÉ O NÍVEL 2		ATÉ O NÍVEL 3		ATÉ O NÍVEL 4		ATÉ O NÍVEL 5		
DE	NO VOCABULÁRIO	CONTROLADO	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%
1	11	0	0,00	11	100,00	-	-	-	-	-	-	-	-	-
2	105	10	9,52	30	28,57	105	100,00	-	-	-	-	-	-	-
3	45	5	11,11	17	37,78	34	75,56	45	100,00	-	-	-	-	-
4	37	1	2,70	13	35,14	27	72,97	35	94,59	37	100,00	-	-	-
5	38	1	2,63	5	13,16	13	34,21	28	73,68	36	94,74	38	100,00	-

Os níveis de subordinação que, cumulativamente, configuraram a maior concentração da defasagem do vocabulário controlado foram o nível 1 (para os documentos do Grupo 1), o nível 2 (para os documentos dos Grupos 2, 3 e 4) e o nível 3 (para os documentos do Grupo 5). Efetivamente, a concentração da

formação de cabeçalhos simples e compostos não disponíveis no vocabulário controlado ocorreu até os níveis 2 e 3 de subordinação dos sumários.

Particularmente no caso dos documentos integrantes dos Grupos 4 e 5, observou-se que, a partir dos últimos níveis de subordinação, a formação de cabeçalhos simples e compostos reduziu, talvez em função da repetição de palavras e/ou de expressões na formação subsequente de cabeçalhos.

Reportando-se ao início do texto do item 8.2, foi verificado que, do total de 677 cabeçalhos simples e compostos formados pela investigação para a Rede BIBLIODATA (CALCO), 236 não estavam incorporados ao vocabulário controlado, o que representou uma defasagem de 34,86% do vocabulário para com a desejável cobertura atualizada da terminologia empregada pelos autores dos livros e monografias examinados (03).

Através do Anexo 29, onde são alfabeticamente listados todos os cabeçalhos de assunto simples e compostos não disponíveis no vocabulário controlado adotado pela Rede, foi possível observar a ocorrência de cabeçalhos comuns a alguns documentos, o que, de certa forma, justificaria a sua incorporação no vocabulário controlado. Entretanto, a maior deficiência do vocabulário verificada durante a formação dos cabeçalhos de assunto foi a impossibilidade de se utilizar algumas palavras sob a forma de subcabeçalhos, as quais seriam indispensáveis para a representação dos conceitos identificados durante a análise conceitual. Um exemplo seria a expressão "INDEXAÇÃO (ECONOMIA)" para a formação de um cabeçalho



representativo do conceito "Indexação de salários", como ocorreu para os documentos 8 e 15. Uma outra limitação do vocabulário refere-se ao impedimento do uso de subcabecçalhos geográficos para determinados cabeçalhos simples, como foi o caso de "COTAS DE IMPORTAÇÃO", o que não permitiu a formação de "COTAS DE IMPORTAÇÃO - BRASIL" para o documento 11.

Também através do Anexo 29 foi verificado que conceitos plenamente estabelecidos no âmbito da literatura sobre Economia não possuíam representatividade no vocabulário da Rede BIBLIODATA (CALCO), nem sob a forma de indicações remissivas, como a exemplo: ANTIDUMPING, ALTA TECNOLOGIA, CRISE FINANCEIRA, CRUZADO (MOEDA), CRUZEIRO (MOEDA), ECONOMIA ABERTA, EMPREGO URBANO, ESTRUTURA INDUSTRIAL, INFLAÇÃO ZERO, POUANÇA ESTATAL, RESERVAS CAMBIAIS e SUBDESENVOLVIMENTO.

Um outro aspecto a abordar seria o fato da Rede BIBLIODATA (CALCO) aceitar a indexação de documentos através de "TERMOS LIVRES", que são cabeçalhos de assunto simples não integrantes do vocabulário controlado mas que podem ser atribuídos pela Biblioteca cooperante interessada, conforme suas necessidades. O uso destes cabeçalhos é exclusivo da Biblioteca que o atribuiu, não sendo, portanto, extensivo às outras Bibliotecas e Centros de Documentação integrantes da Rede. A disponibilidade deste recurso na Rede BIBLIODATA (CALCO) não impediria ou limitaria a atribuição de cabeçalhos de assunto simples, como os acima identificados e pertinentes para a representação conceitual dos livros e monografias investigados. A menção deste recurso da Rede na presente Dissertação objetivou atentar para o fato de que existe a possibilidade de se conferir um número maior de

cabeçalhos aos documentos independentemente das restrições impostas pelo vocabulário controlado adotado.

#### 6.2.3.1.2 Os cabeçalhos simples e compostos coincidentes.

Este item refere-se aos cabeçalhos formados a partir da análise conceitual mas identificados como coincidentes com alguns daqueles atribuídos aos livros e monografias integrantes da amostra durante a catalogação por assunto executada pela Rede BIBLIODATA (CALCO) e que, por conseqüência, não foram incorporados ao grupo de cabeçalhos acrescidos pela investigação.

Conforme os dados quantitativos incluídos no Anexo 21, os cabeçalhos simples e compostos coincidentes foram computados para cada Grupo de documentos, de acordo com os níveis de subordinação onde foram formados e identificados, como apresentado no Quadro 44, a seguir:

Quadro 44: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CADEÇALHOS SIMPLES E COMPOSTOS COINCIDENTES IDENTIFICADOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO					TOTALS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	
1	0	3	-	-	-	3
2	6	6	5	-	-	17
3	4	3	0	1	-	8
4	0	0	0	0	1	1
5	0	0	0	0	0	0

Todos os cabeçalhos simples e compostos coincidentes acima quantificados (que receberam, respectivamente, as codificações A1 e C1, como consta no item 7.2.3) são discriminados no Anexo 30, para cada documento e por nível de subordinação onde foram formados.

Através do Quadro 44, observou-se que poucos foram os cabeçalhos simples e compostos distinguidos como coincidentes com a catalogação por assunto efetuada pela Rede BIBLIODATA (CALCO) (29 cabeçalhos para todos os documentos integrantes da amostra). Comparando-se este valor com o total de cabeçalhos simples e compostos conferidos pela Rede aos 30 livros e

monografias (ver Anexo 23), que corresponde a 81 cabeçalhos, verifica-se que os cabeçalhos coincidentes responderam por 35,80% deste total. Observou-se também que a maior parte dos cabeçalhos foi identificada nos primeiros níveis de subordinação, sem, entretanto, poder ficar caracterizada uma tendência de utilização dos títulos, subtítulos e sumários, pelos profissionais da Rede, como fontes de palavras para a representação conceitual do conteúdo dos documentos.

#### B.2.3.1.3 - A demonstração do potencial na Rede BIBLIODATA (CALCO).

Para se demonstrar e analisar o potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões passíveis de representação do conteúdo informativo veiculado por livros e monografias em Sistemas de Informação pré-coordenados, foi necessário quantificar, conjuntamente, todos os cabeçalhos simples e compostos acrescidos, não disponíveis no vocabulário controlado e coincidentes anterior e respectivamente quantificados através dos Quadros 22, 41 e 44. O Quadro 45, a seguir, apresenta o somatório destes cabeçalhos para cada Grupo de documentos, de acordo com os níveis de subordinação onde foram formados.

Quadro 45: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS  
 CADEÇALHOS SIMPLES E COMPOSTOS ACRESCIDOS, NÃO  
 DISPONÍVEIS NO VOC. CONTROLADO E COINCIDENTES,  
 POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS
1	3	37	-	-	-	-	40
2	31	72	204	-	-	-	307
3	13	26	52	24	-	-	115
4	3	24	23	17	9	-	76
5	3	9	42	50	29	6	139

O total apontado de 677 cabeçalhos formados correspondeu a:

- 412 cabeçalhos simples e compostos acrescidos (60,86%),
- 236 cabeçalhos simples e compostos não disponíveis no vocabulário controlado (34,86%), e
- 29 cabeçalhos simples e compostos coincidentes (4,28%).

Algumas análises foram elaboradas através das quantificações indicadas no Quadro 45.

Calculando-se a média de cabeçalhos formados a partir dos títulos, subtítulos e sumários para cada documento, foi possível observar que quanto mais hierarquizadas eram as estruturas dos sumários, maiores foram as possibilidades de formação de cabeçalhos, mesmo considerando-se o discretíssimo decréscimo

ocorrido nas médias dos documentos do Grupo 2 para o Grupo 3, como apresentado no Quadro 46, a seguir:

Quadro 46: REDE BIBLIODATA (CALCO): MÉDIAS DO TOTAL DE CADEÇALHOS SIMPLES E COMPOSTOS FORMADOS, POR DOCUMENTO.

GRUPOS	NÚMERO DE DOCUMENTOS NO GRUPO	TOTAL DE CADEÇALHOS SIMPLES E COMPOSTOS FORMADOS NO GRUPO	MÉDIA DE CADEÇALHOS POR DOCUMENTO
1	4	40	10,00
2	16	307	19,19
3	6	115	19,17
4	2	76	38,00
5	2	139	69,50

Para a análise da participação de cada nível de subordinação no total de cabeçalhos formados para os Grupos de documentos, foi necessário efetuar os respectivos cálculos percentuais, como indicado no Quadro 47, a seguir:

Quadro 47: REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS DE ASSUNTO SIMPLES E COMPOSTOS FORMADOS, POR GRUPOS DE DOCUMENTOS.

GRUPO	VALORES	VALORES APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO											
		TOTAIS											
DE	APURADOS	NÍVEL 00		NÍVEL 01		NÍVEL 02		NÍVEL 03		NÍVEL 04		NÍVEL 05	
		NO											
DOCUMENTOS	GRUPO	OBtido:	%	OBtido:	%	OBtido:	%	OBtido:	%	OBtido:	%	OBtido:	%
		1	40	3	7,50	37	92,50	-	-	-	-	-	-
2	307	31	10,10	72	23,45	204	66,45	-	-	-	-	-	-
3	115	13	11,30	26	22,61	52	45,22	24	20,87	-	-	-	-
4	76	3	3,95	24	31,58	23	30,26	17	22,37	9	11,84	-	-
5	139	3	2,16	9	6,48	42	30,22	50	35,97	29	20,85	6	4,32

Por intermédio do Quadro 47, verificou-se que, para os Grupos 2, 3 e 4, a maior parte da formação de cabeçalhos ocorreu nos níveis 1 e 2 de subordinação, decaindo nos níveis subseqüentes. Para o Grupo 5, a formação de cabeçalhos foi mais dispersiva entre os níveis de subordinação, centralizando-se nos níveis 2 e 3 e também decaindo a partir dos níveis subseqüentes. No Grupo 1, a formação de cabeçalhos foi obtida praticamente no nível 1 de subordinação, fato justificado pela pequena especificidade da abordagem dos assuntos tratados pelos documentos demonstrada pela estrutura de seus sumários.

Outra observação refere-se à contribuição dos títulos e subtítulos na formação de cabeçalhos (nível 0 de subordinação). Comparando-a com as quantificações e percentuais apurados nos níveis de subordinação correspondentes aos sumários, ela foi sensivelmente inferior.

Dando prosseguimento à análise da formação de cabeçalhos, foi efetuada a soma dos percentuais apurados no Quadro 47, na seqüência dos níveis de subordinação. Observou-se que, para os documentos integrantes dos Grupos 1, 2 e 3, obteve-se a formação de, no mínimo, 79,13% dos cabeçalhos reccorendo-se cumulativamente até o nível 2 de subordinação dos seus sumários. Para os livros e monografias integrantes dos Grupos 4 e 5, foi constatado que, examinando-se cumulativamente até o nível 3 de subordinação de seus sumários, alcançou-se a formação de, no mínimo, 74,83% dos cabeçalhos. Estes dados podem ser confirmados através do Quadro 48, a seguir apresentado, o qual inclui a soma cumulativa dos percentuais apurados no Quadro 47, na seqüência dos níveis de subordinação, por Grupos de documentos.



Quadro 48: REDE BIBLIODATA (CALCO): PERCENTUAIS DE CABEÇALHOS DE ASSUNTO FORMADOS, ACUMULADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO				
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
1	7,50	100,00	-	-	-
2	10,10	33,55	100,00	-	-
3	11,30	33,91	79,13	100,00	-
4	3,95	35,53	65,79	88,16	100,00
5	2,16	8,64	38,86	74,83	95,68

Com base em todas as quantificações e análises referentes aos cabeçalhos de assunto simples e compostos formados para a Rede BIBLIODATA (CALCO) a partir da análise conceitual executada através dos títulos, subtítulos e sumários dos 30 livros e monografias integrantes da amostra, ficou demonstrada a possibilidade de se atribuir um considerável número de cabeçalhos de assunto destinados à recuperação de informações via pré-coordenação.

Também ficou demonstrado que os sumários de livros e monografias, junto com os seus títulos e subtítulos, realmente são elementos estruturais capazes de proporcionar indicações bem pertinentes à representação do conteúdo informativo destes documentos, sem requererem dos profissionais da informação uma interpretação de conteúdo tão profunda, ainda mais no caso da pré-coordenação, a qual ao determinar uma contextualização para a representação conceitual, seria bem atendida quanto à formação e à atribuição de cabeçalhos, uma vez que entende-se a estrutura dos sumários, dada a sua hierarquização, como naturalmente contextualizada.

A demonstração do potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos, assim como as quantificações apuradas para todos os cabeçalhos de assunto formados e os respectivos cálculos percentuais permitiram inferir que, recorrer até o nível 2 de subordinação dos sumários para livros e monografias com menor especificidade de abordagem de seus temas centrais e recorrer até o nível 3 de subordinação dos sumários para livros e monografias com maior especificidade, parece ser suficiente para se alcançar uma maior e mais proveitosa representação do conteúdo informativo destes documentos em Sistemas de Informação pré-coordenados, e, conseqüentemente, provocar uma maior disponibilidade de meios de acesso a documentos tão estabelecidos na literatura especializada. Estas constatações já foram expressadas quando da quantificação e análise dos acréscimos obtidos pela investigação junto à Rede BIBLIODATA (CALCO), no item 8.2.1, o que pode corroborá-las.

é sabido que uma diretriz relativa à utilização conjunta de títulos, subtítulos e sumários de livros e monografias para a identificação e posterior representação conceitual ficaria condicionada à Política de Indexação vigente para o Sistema de Informação, a qual, ao definir os padrões de exaustividade e de especificidade de indexação dos documentos, recomendaria recorrer-se ao nível de subordinação dos sumários que, como demonstrado pela presente Dissertação, fosse compatível com os referidos padrões.

Reportando-se especialmente à Rede BIBLIODATA (CALCO), um Sistema de Informação cooperativo, com atuação a nível nacional, cujas Bibliotecas e Centros de Documentação integrantes, além de possuírem acervos valiosíssimos (como é o caso da Fundação Biblioteca Nacional), atendem a categorias bem distintas de usuários, cabe sugerir uma revisão em sua Política de Indexação no sentido de estabelecer normas e mecanismos que ocasionassem uma maior disponibilidade de meios de acesso ao conteúdo informativo de livros e monografias a processar. Esta sugestão, além dos expressivos resultados obtidos pela presente investigação em Sistemas de Informação pré-coordenados, baseou-se também no reduzido número de cabeçalhos de assunto conferidos aos 30 documentos integrantes da amostra examinada pelas Bibliotecas e Centros de Documentação componentes da Rede, fato que, certamente, se opõe aos objetivos fim definidos para o estabelecimento e operacionalização da Rede BIBLIODATA (CALCO).

Talvez, como já manifestado durante a análise dos resultados relativos à pós-coordenação para a Biblioteca do BNDES, o fator "cultural" da indexação superficial de livros e monografias, por não serem documentos com informações de caráter imediatista, também pode ter sido uma das razões que contribuíram para a atribuição de poucos cabeçalhos de assunto pela Rede, ou então diretrizes dos administradores das Bibliotecas e Centros de Documentação cooperantes quanto à redução de custos operacionais, ou até mesmo a falta de recursos humanos.

A atribuição de um maior número de cabeçalhos de assunto para livros e monografias pode ser até questionada em função dos princípios de exaustividade e de especificidade de indexação de documentos. O ideal seria atingir um equilíbrio entre estes dois postulados, mas eles, na teoria e na prática, são vinculados a uma Política de Indexação definida em função dos objetivos, do material bibliográfico a processar e, principalmente, dos usuários dos Sistemas de Informação.

Entretanto, é primordial observar que os resultados apontados para a Rede BIBLIODATA (CALCO) demonstraram a flexibilidade de aplicação da metodologia utilizada durante a pesquisa para o alcance de níveis de exaustividade ou de precisão estipulados em diretrizes ou Políticas de Indexação fixadas pelos Sistemas de Informação em relação a livros e monografias.

Além disso, os resultados apurados junto à Rede BIBLIODATA (CALCO) demonstraram, principalmente, a exequibilidade de um novo método destinado ao aprimoramento da catalogação por assunto de livros e monografias, e, ao mesmo tempo, configuraram a necessidade de se alterar o atual posicionamento técnico das Bibliotecas e Centros de Documentação cooperantes quanto à superficialidade da indexação destas categorias de documentos.

#### 8.2.3.2 - O potencial na Base de Dados bibliográficos da Rede BIBLIODATA (CALCO).

Para se efetuar a demonstração do potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões representativas de conceitos veiculados por livros e monografias junto à Base de Dados bibliográficos da Rede BIBLIODATA (CALCO), foi necessário quantificar e analisar, além das variáveis já examinadas no item 8.2.2.2, os cabeçalhos tópicos e os subcabeçalhos que compuseram a formação dos cabeçalhos compostos não disponíveis no vocabulário controlado e a formação dos cabeçalhos compostos coincidentes.

Antes, porém, foi necessário esclarecer que, pelo fato dos subcabeçalhos cronológicos atuarem na Base de Dados bibliográficos somente como parâmetro de restrição de busca, a demonstração do potencial foi efetuada de duas formas: excluindo-se e incluindo-se estes subcabeçalhos.

Dando início à análise das variáveis complementares, a partir do Anexo 20, os cabeçalhos tópicos e os subcabeçalhos referentes aos cabeçalhos compostos não disponíveis no vocabulário controlado adotado pela Rede BIBLIODATA (CALCO) foram quantificados para cada Grupo de documentos, de acordo com os níveis de subordinação, como apresentado no Quadro 49, a seguir:

Quadro 49: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES DOS CABEÇALHOS COMPOSTOS NÃO DISPONÍVEIS NO VOCABULÁRIO CONTROLADO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO:
	NÍVEL 0   NÍVEL 1   NÍVEL 2   NÍVEL 3   NÍVEL 4   NÍVEL 5   TOTAIS
1	0   10   -   -   -   -   10
2	6   11   38   -   -   -   55
3	4   8   10   2   -   -   24
4	1   0   8   3   1   -   21
5	2   1   3   6   2   1   15

Todos os cabeçalhos tópicos e subcabeçalhos acima quantificados (que receberam, respectivamente, as codificações E1 e E2, como consta no item 7.3.2.1) encontram-se discriminados no Anexo 31, por documento e incluindo os níveis de subordinação onde foram formados.

Retomando a deficiência do vocabulário controlado adotado pela Rede, já analisada no item 8.2.3.1.1, através do Anexo 31 foi possível verificar que alguns conceitos bem estabelecidos no campo Economia não estavam incluídos no vocabulário, sob a forma de cabeçalho tópico ou de subcabeçalho. Dentre eles, "ANÁLISE ECONÔMICA" (como subcabeçalho nos documentos 06 e 24), "CRISE FINANCEIRA" (como subcabeçalho nos documentos 21 e 23), "INDEXAÇÃO (ECONOMIA)" (como subcabeçalho nos documentos 08 e 18), "INVESTIMENTOS" (como subcabeçalho nos documentos 21 e 30), e "RISCOS (ECONOMIA)" (como subcabeçalho nos documentos 06 e 18).

Continuando a quantificação das variáveis consideradas para a demonstração do potencial dos títulos, subtítulos e sumários para a Base de Dados bibliográficos, os cabeçalhos tópicos e subcabeçalhos identificados pela investigação como coincidentes com alguns daqueles que compuseram a formação dos cabeçalhos compostos atribuídos durante a catalogação por assunto executada pela Rede BIBLIODATA (CALCO), foram computados para cada um dos Grupos de documentos, conforme apresentado no Quadro 50, a seguir:

Quadro 50: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DOS  
 CABEÇALHOS TÓPICOS E SUBCABEÇALHOS INTEGRANTES  
 DOS CABEÇALHOS COMPOSTOS COINCIDENTES, POR  
 GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS:	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS
1	3	5	-	-	-	-	8
2	16	9	3	-	-	-	28
3	7	3	2	1	-	-	13
4	3	1	0	0	0	-	4
5	2	0	0	0	0	0	2

Em relação aos subcabecçalhos cronológicos coincidentes, apenas 2(dois) foram identificados pela investigação: um para o grupo 2 de documentos, no nível 1 de subordinação, e outro para o Grupo 5 de documentos, também no nível 1 de subordinação (ver Anexo 21).

Todos os cabeçalhos tópicos, subcabecçalhos e subcabecçalhos cronológicos identificados como coincidentes (que receberam, respectivamente, as codificações B4, B5 e B6, como consta no item 7.3.2.1) encontram-se discriminados no Anexo 32, por documento, incluindo os níveis de subordinação onde foram identificados.



Foi observado que a maior incidência de cabeçalhos tópicos, subcabeçalhos e subcabeçalhos cronológicos coincidentes ocorreu nos níveis 0 e 1 de subordinação, sendo que, nos demais níveis praticamente não houve a identificação destas variáveis.

Após a quantificação dos cabeçalhos tópicos, subcabeçalhos e subcabeçalhos cronológicos não disponíveis no vocabulário controlado e coincidentes, foi possível dar início à demonstração propriamente dita do potencial dos títulos, subtítulos e sumários para a Base de Dados bibliográficos da Rede BIBLIODATA (CALCO).

#### 8.2.3.2.1 -- A demonstração do potencial sem os subcabeçalhos cronológicos.

A demonstração deste potencial requereu a quantificação conjunta das seguintes variáveis (a inclusão do número dos Quadros correspondentes às variáveis discriminadas objetivou facilitar o esclarecimento dos valores considerados):

##### A) os cabeçalhos simples e compostos:

- . acrescidos (Quadro 22);
- . não disponíveis no vocabulário controlado (Quadro 41);
- . coincidentes (Quadro 44); e

B) os cabeçalhos tópicos e subcabeçalhos que compuseram a formação dos cabeçalhos compostos:

. acrescidos (Quadro 30);

. não disponíveis no vocabulário controlado (Quadro 49);

. coincidentes (Quadro 50).

Tomando-se os valores incluídos nos quadros acima referenciados, o somatório de todos estes cabeçalhos e subcabeçalhos é apresentado no Quadro 51, a seguir, por Grupo de documentos e de acordo com o nível de subordinação onde foram formados ou identificados.

Quadro 51: REDE BIBLIODATA (CALCO): QUANTIFICAÇÃO DO POTENCIAL DOS TÍTULOS, SUBTÍTULOS E SUMÁRIOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, SEM OS SUBCABEÇALHOS CRONOLÓGICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	VALORES TOTAIS APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO						TOTAIS
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	
1	9	61	-	-	-	-	70
2	44	124	315	-	-	-	503
3	28	43	81	38	-	-	190
4	7	44	36	23	11	-	121
5	9	12	67	69	35	8	200

Algumas análises foram executadas a partir das quantificações indicadas no Quadro 51 acima.

Calculando-se a média de cabeçalhos e subcabeçalhos, formados ou identificados pela investigação através dos títulos, subtítulos e sumários, por documento, foi verificado que estas médias foram se elevando à medida que a especificidade da abordagem dos temas centrais dos livros e monografias era maior. Observa-se por meio do Quadro 52, a seguir, que as médias quase que dobraram de valor do Grupo 1 para o Grupo 2, como também do Grupo 3 para o Grupo 4, e do Grupo 4 para o Grupo 5:

Quadro 52: REDE BIBLIODATA (CALCO): MÉDIAS DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR DOCUMENTO.

GRUPOS DE DOCUMENTOS	NÚMERO DE DOCUMENTOS NO GRUPO	TOTAL DE PONTOS DE ACESSO FORMADOS PARA O GRUPO	MÉDIA DE CABEÇALHOS POR DOCUMENTO
1	4	70	17,50
2	16	503	31,44
3	6	190	31,67
4	2	121	60,50
5	2	200	100,00

A análise da contribuição percentual de cada nível de subordinação em relação ao total de cabeçalhos e subcabeçalhos formados em cada Grupo de documentos indicou que:

- para os Grupos 2 e 3 de documentos, o nível 2 de subordinação foi aquele que possibilitou a formação da maior parte dos cabeçalhos e subcabeçalhos;
- no Grupo 4 de documentos, a maior formação de cabeçalhos e subcabeçalhos ocorreu no nível 1 de subordinação (36,36%); o nível 2 de subordinação participou com valores próximos dos apontados no nível 1 (29,75%);
- no Grupo 5 de documentos, o nível 3 de subordinação proporcionou a formação da maior parte dos cabeçalhos e subcabeçalhos (34,50%) e o nível 2 contribuiu com valores também muito próximos (33,50%);
- para o Grupo 1 de documentos, devido à disponibilidade de um único nível de subordinação, praticamente todos os cabeçalhos e subcabeçalhos foram formados no nível 1 de subordinação (81,14%);
- de um modo geral, os níveis 1 e 2 de subordinação foram aqueles que permitiram a formação da maior parte dos cabeçalhos e subcabeçalhos, com exceção do Grupo 5 de documentos, cuja formação da maior parte de cabeçalhos e subcabeçalhos ocorreu nos níveis 2 e 3 de subordinação;

em todos os Grupos de documentos, os títulos e subtítulos (nível 0 de subordinação) tiveram uma participação muito pequena na formação dos cabeçalhos e subcabeçalhos, comparando-os com os sumários; e

no caso dos documentos integrantes dos Grupos 3, 4, e 5, a formação de cabeçalhos e subcabeçalhos decaiu nos últimos níveis de subordinação.

Estas constatações podem ser confirmadas através do Quadro 53, a seguir apresentado, o qual inclui os cálculos percentuais da contribuição de cada nível de subordinação em relação ao total de cabeçalhos e subcabeçalhos formados para os Grupos de documentos constituídos:

Quadro 53: REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS	TOTALS DE	VALORES APURADOS EM CADA NÍVEL DE SUBORDINAÇÃO													
		CABEÇALHOS		NÍVEL 0		NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4		NÍVEL 5	
DE	E														
		SUBCABEÇALHOS													
DOCUMENTOS	FORMADOS	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%	OBTIDO	%
1	70	9	12,86	61	87,14	-	-	-	-	-	-	-	-	-	-
2	503	64	12,72	124	24,65	315	62,63	-	-	-	-	-	-	-	-
3	190	28	14,74	43	22,63	81	42,63	38	20,00	-	-	-	-	-	-
4	121	7	5,79	44	36,36	36	29,75	23	19,01	11	9,09	-	-	-	-
5	200	9	4,50	12	6,00	67	33,50	69	34,50	35	17,50	8	4,00	-	-

Foi também oportuno avaliar a soma cumulativa dos percentuais incluídos no Quadro 53 acima, na seqüência dos níveis de subordinação. As somas mencionadas indicaram que, recorrendo-se até o nível 2 de subordinação, foi possível formar, no mínimo, 71,90% dos cabeçalhos e subcabeçalhos. Esta constatação não pôde ser estendida ao Grupo 5 de documentos o qual, somente a partir do nível 3 de subordinação, obteve 78,50% da formação de cabeçalhos e subcabeçalhos. O Quadro 54, a seguir, pode comprovar estes resultados:

Quadro 54: REDE BIBLIODATA (CALCO): PERCENTUAIS DE CADEÇALHOS E SUBCADEÇALHOS FORMADOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO				
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
1	12,06	100,00	-	-	-
2	12,72	37,37	100,00	-	-
3	14,74	37,37	60,00	100,00	-
4	5,79	42,15	71,90	90,91	100,00
5	4,50	10,50	44,00	78,50	96,00

Com base em todos os valores apurados e nas respectivas análises efetuadas relativas à demonstração do potencial dos títulos, subtítulos e sumários como fontes de palavras e/ou expressões destinadas à formação de cabeçalhos de assunto, foi possível inferir que, para a Base de Dados bibliográficos da Rede BIBLIODATA (CALCO), recorrer pelo menos até o nível 2 de subordinação dos sumários (não se omitindo os títulos e os subtítulos) é extremamente proveitoso no sentido de tornar o conteúdo informativo específico de livros e monografias mais acessível pelos recursos operacionais da referida Base e, concomitantemente, executar-se uma indexação consistente e

condizente com a riqueza das informações veiculadas por essas categorias de documentos.

Os resultados do potencial apontados para a Base de Dados foram muito semelhantes àqueles descritos para a Rede BIBLIODATA (CALCO) no item 8.2.3.1, os quais também indicaram a utilização até o nível 2 de subordinação dos sumários para a identificação de palavras e/ou expressões com elevado grau de representatividade conceitual.

#### 8.2.3.2.2 - A demonstração do potencial com os subcabecçalhos cronológicos.

Pelo fato dos subcabecçalhos cronológicos atuarem na Base de Dados bibliográficos da Rede BIBLIODATA (CALCO) somente como parâmetro restritivo na formulação de estratégias de busca, para se efetuar a demonstração do potencial dos títulos, subtítulos e sumários considerando-se esta característica operacional do Sistema, foi necessário quantificar, conjuntamente, os valores incluídos no Quadro 51 do item 8.2.3.2.1, os valores referentes aos subcabecçalhos cronológicos acrescidos pela investigação (já computados através do Quadro 34 e analisados no item 8.2.2.2) e os 2(dois) subcabecçalhos cronológicos coincidentes identificados pela investigação e já discriminados (ver p. 178). O Quadro 55, a seguir, apresenta o somatório de todas estas variáveis, por Grupo de documentos, e de acordo com o nível de subordinação onde foram formados ou identificados:



Quadro 55: REDE BIBLIODATA - (CALCO): QUANTIFICAÇÃO DO POTENCIAL DOS TÍTULOS, SUBTÍTULOS E SUMÁRIOS PARA A BASE DE DADOS BIBLIOGRÁFICOS, COM OS SUBCABEÇALHOS CRONOLÓGICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	TOTAIS
1	10	68	-	-	-	-	78
2	65	132	332	-	-	-	529
3	28	44	82	41	-	-	195
4	7	47	39	30	11	-	134
5	9	16	72	76	37	0	210

Efetuada-se uma análise baseada nos valores acima indicados e semelhante à executada para a demonstração do potencial dos títulos, subtítulos e sumários sem os subcabecçalhos cronológicos, foi calculada a contribuição percentual de cada nível de subordinação no total de cabeçalhos e subcabecçalhos formados. Os resultados apurados foram praticamente idênticos àqueles relativos ao potencial sem se considerar os subcabecçalhos cronológicos (ver Quadro 53), os quais revelaram a formação da maior parte dos cabeçalhos nos níveis 1 e 2 de subordinação para os Grupos 1, 2, 3 e 4 de documentos, e nos níveis 2 e 3 para o Grupo 5 de documentos. O

Quadro 56, a seguir, pode confirmar estas constatações.

Quadro 56: REDE BIBLIODATA (CALCO): CONTRIBUIÇÃO DE CADA NÍVEL DE SUBORDINAÇÃO EM RELAÇÃO AO TOTAL DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS (INCLUINDO OS CRONOLÓGICOS), PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	TOTALS DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS	VALORES AFURADOS EM CADA NÍVEL DE SUBORDINAÇÃO									
		NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5	OBTIDO	%	OBTIDO	%
1	78	10	68	-	-	-	-	10	12,82	68	87,18
2	529	65	332	-	-	-	-	65	12,29	332	62,76
3	195	28	82	41	-	-	-	28	14,36	82	42,05
4	134	7	39	30	11	-	-	7	5,22	47	35,08
5	218	9	72	76	37	8	-	9	4,13	16	7,34

Continuando a análise, foi efetuada a soma cumulativa dos percentuais acima indicados, na seqüência dos níveis de subordinação, a qual também apontou resultados muito aproximados daqueles relativos ao potencial sem os subcabçalhos cronológicos (ver Quadro 54). O Quadro 57, a seguir, apresenta esta soma cumulativa, por Grupos de documentos e de acordo com os níveis de subordinação.

Quadro 57: REDE BIBLIODATA. (CALCO): PERCENTUAIS DE CABEÇALHOS E SUBCABEÇALHOS FORMADOS, APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO (INCLUINDO OS SUBCABEÇALHOS CRONOLÓGICOS), PARA A BASE DE DADOS BIBLIOGRÁFICOS, POR GRUPOS DE DOCUMENTOS.

GRUPOS DE DOCUMENTOS	PERCENTUAIS APURADOS NA SEQUÊNCIA DOS NÍVEIS DE SUBORDINAÇÃO				
	NÍVEL 0	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
1	12,82	100,00	-	-	-
2	12,29	37,24	100,00	-	-
3	14,36	36,92	70,97	100,00	-
4	5,22	40,30	69,40	91,79	100,00
5	4,13	11,47	44,50	79,36	96,93

Os valores acima indicados também revelaram que cerca de 70,00% dos cabeçalhos e subcabeçalhos foram formados até o nível 2 de subordinação dos sumários, para os documentos integrantes dos Grupos 1, 2, 3 e 4, e cerca de 80,00% até o nível 3 de subordinação para os documentos do Grupo 5.

Os resultados do potencial dos títulos, subtítulos e sumários para a Base de Dados bibliográficos, considerando-se os subcabeçalhos cronológicos, vieram corroborar aqueles apontados quando se excetuou esta categoria de subcabeçalhos, confirmando, assim, a constatação de que, recorrer-se até o nível 2 de subordinação dos sumários de livros e monografias com menor especificidade de abordagem de seus temas centrais e até o nível 3 para livros e monografias com maior especificidade, é exequível e muito proveitoso para a identificação de palavras e/ou expressões com considerável grau de representatividade da sobrecedência completa desses documentos.

Cabe reiterar que a aplicação da técnica proposta pela presente Dissertação ficaria condicionada à Política de Indexação vigente nas Bibliotecas e Centros de Informação integrantes da Rede BIBLIODATA (CALCO), a qual, ao indicar o nível de especificidade da catalogação por assunto, poderia mencionar o nível de subordinação dos sumários compatível com a referida especificidade para a identificação de palavras e/ou expressões potencialmente representativas de conceitos.

Retomando as características operacionais da Base de Dados bibliográficos, é oportuno mencionar que a administração da Rede BIBLIODATA (CALCO) encontra-se efetuando pesquisas e estudos no sentido de desenvolver os mecanismos de recuperação de informações (04). Um deles refere-se à possibilidade de intersecção de palavras durante a formulação da estratégia de busca por assunto. Se isto ocorrer, o aumento do número de cabeçalhos de assunto atribuídos por livro ou monografia, de

acordo com a proposta da presente Dissertação, certamente tornaria a Base de Dados muito mais rica para responder eficazmente às solicitações dos usuários e, paralelamente, provocaria uma maior disponibilidade de acesso ao conteúdo específico desses documentos.

De um modo geral, os resultados da investigação relativos a Sistemas de Informação que adotem a pré-coordenação como técnica de indexação e de recuperação de informações foram muito favoráveis. Tanto os acréscimos obtidos quanto a demonstração do potencial dos títulos, subtítulos e sumários foram extremamente reveladoras.

### 0.3 - Comparação dos resultados obtidos para a pós-coordenação e para a pré-coordenação.

Como já mencionado, em ambos os Sistemas de Informação examinados, os resultados obtidos pela presente investigação foram surpreendentes, tanto no que se refere à apuração dos acréscimos quanto à demonstração do potencial que os títulos, subtítulos e sumários de livros e monografias possuem, conjuntamente, para a identificação de palavras, expressões e nomes próprios, com elevado índice de representatividade conceitual, destinados à recuperação das informações que estes documentos veiculam.

Para os dois Sistemas de Informação investigados, ficou evidenciado pelos resultados apontados que, cumulativamente, os níveis 1, 2 e 3 de subordinação dos sumários estabelecidos pela metodologia da presente Dissertação são aqueles que oferecem a maior cobertura consistente das abrangências geral e específica do conteúdo informativo de livros e monografias. O nível 3 de subordinação compreende os tópicos mais específicos dos trabalhos, e deve ser utilizado em indexações destinadas à obtenção de elevados níveis de precisão na recuperação de informações. A partir dos níveis 4 e 5 de subordinação dos sumários, o número de palavras e/ou expressões representativas diminuiu em relação ao que foi obtido nos níveis anteriores, devido a repetições e/ou a intitulação de tópicos dos sumários com pequeno número de palavras inéditas para o contexto dos documentos.

Lamentavelmente, foi demonstrado que os dois Sistemas de Informação examinados atribuíram pouquíssimos cabeçalhos de assunto para os 30 documentos integrantes da amostra (média de 4,87 cabeçalhos e associações por documento na Biblioteca do BNDES, e de 2,70 cabeçalhos de assunto simples e compostos por documento na Rede BIBLIODATA (CALCO)), o que representou uma elevada limitação de acesso ao conteúdo informativo completo dos livros e monografias. Este fator ocasionou a obtenção de acréscimos tão expressivos, principalmente para os documentos dos Grupos 4 e 5, os quais possuíam sumários que denotavam uma abordagem bem específica dos temas centrais dos trabalhos e que, no entanto, possuíam um reduzidíssimo número de cabeçalhos de assunto atribuídos pelos Sistemas para posterior recuperação de sua sobrecedência completa.

Os resultados, também e oportunamente, revelaram a desatualização e a deficiência dos vocabulários controlados utilizados pelos dois Sistemas em relação à desejável cobertura da terminologia utilizada pela literatura sobre Economia. Na Rede BIBLIODATA (CALCO), quando examinado o potencial dos títulos, subtítulos e sumários para a formação de cabeçalhos de assunto, verificou-se que o vocabulário controlado estava 34,86% defasado quanto às palavras e/ou expressões identificadas como relevantes para a representação conceitual dos documentos integrantes da amostra. Na Biblioteca do BNDES este percentual alcançou 42,11%, fato que surpreendentemente não deveria ocorrer dada a liberdade de atribuição de cabeçalhos vigente naquele Sistema, como também devido à especialização do vocabulário ser voltada para o campo Economia.

Ficou comprovado que os títulos e subtítulos, sozinhos, não oferecem um número de palavras suficientemente adequado para uma representação consistente do conteúdo informativo de livros e monografias. Para os dois Sistemas de Informação investigados, os títulos e subtítulos, em quaisquer das diversas apurações efetuadas, não conseguiram fornecer 15% das palavras e/ou expressões identificadas no total. Com a participação conjunta dos sumários na identificação de palavras, foi possível obter resultados tão expressivos.

Um outro importante aspecto a ressaltar refere-se à disponibilidade de automação de serviços existente nos dois Sistemas de Informação examinados, principalmente aqueles destinados aos processos de indexação e de recuperação de

informações. Tanto a Biblioteca do BNDES quanto a Rede BIBLIODATA (CALCO), dada a capacidade operacional dos "software" que utilizam, certamente poderiam absorver um volume bem maior de entradas por assunto.

O decorrente aumento do custo operacional dos Sistemas, certamente, seria compensado pela otimização do uso de seus acervos bibliográficos, a qual seria manifestada através do fornecimento aos usuários de um maior número de respostas precisas e consistentes na recuperação de informações veiculadas pelas coleções de livros e monografias.

Concluindo, cabe mencionar que os resultados obtidos junto aos dois Sistemas de Informação examinados foram homogêneos, demonstrando o equilíbrio de exeqüibilidade da técnica metodológica igualmente empregada em Sistemas com características tão distintas, principalmente aquelas referentes aos recursos de indexação e de recuperação de informações que utilizam.

Ainda concluindo a avaliação dos resultados auferidos tanto para a pós quanto para a pré-coordenação, cabe informar que, por motivos de ordem operacional dos Sistemas de Informação examinados (05), infelizmente não foi possível efetuar um teste destinado à mensuração dos índices de revocação, de relevância e de precisão a ocorrerem em recuperações de informações nos livros e monografias integrantes da amostra examinada, indexados conforme a metodologia aplicada pela presente investigação. As revelações desse teste certamente complementariam e



enriqueceriam sobremaneira a discussão dos resultados alcançados já descritos, e, principalmente, confeririam um maior respaldo metodológico para a Dissertação.

#### 8.4 - Relações entre a quantificação dos cabeçalhos de assunto formados e a quantificação das páginas de texto correspondentes aos níveis de subordinação dos sumários dos documentos integrantes da amostra examinada

Como previsto pela metodologia aplicada na presente Dissertação, para complementar a discussão dos resultados anteriormente apresentados, foi executada uma mensuração da abrangência de contextualização do conteúdo informativo dos livros e monografias proporcionada pelas páginas de texto correspondentes aos níveis de subordinação definidos para os sumários dos documentos examinados.

Este estudo objetivou identificar e avaliar prováveis relações existentes entre o número de cabeçalhos quantificados durante as demonstrações do potencial dos títulos, subtítulos e sumários como fontes de palavras representativas de conceitos veiculados em livros e monografias para a indexação e recuperação de informações em Sistemas de Informação pós e pré-coordenados (itens 8.1.2 e 8.2.3.1), e o número de páginas de texto compreendidas pelos níveis de subordinação onde foram computados os cabeçalhos de assunto formados a partir das palavras representativas identificadas durante a análise conceitual.

Foi efetuada a contagem das páginas de texto referentes a cada nível de subordinação estabelecido para os sumários dos documentos integrantes da amostra, assim como foram quantificadas as incidências de níveis de subordinação nos sumários. As páginas de texto computadas em cada nível de subordinação foram percentualmente comparadas com o total de páginas de texto dos documentos.

Foram excluídos desta avaliação os documentos integrantes do Grupo 1, os quais, por possuírem apenas um único nível de subordinação em seus sumários, obviamente não foi necessário examinar. Portanto, a apreciação foi executada nos sumários dos documentos integrantes dos Grupos 2, 3, 4 e 5.

Nestes Grupos de documentos examinados, o nível 1 de subordinação não foi considerado na contagem de páginas de texto, uma vez que ele, por natureza, compreenderia 100,00% da abrangência dos tópicos abordados pelos livros e monografias.

O Anexo 33 apresenta as quantificações relativas às incidências de níveis de subordinação nos sumários e às páginas de texto correspondentes aos níveis considerados, como também os cálculos percentuais das páginas de texto quantificadas nos níveis de subordinação, em relação ao total de páginas de texto dos livros e monografias, para cada um dos Grupos de documentos examinados.

Reportando-se ao Anexo 33, as médias dos percentuais de páginas de texto ocorridos nos níveis de subordinação em relação ao total de páginas de texto dos documentos integrantes dos Grupos examinados foram como a seguir discriminadas:

A) Grupo 2 de documentos:

. no nível 2 de subordinação: média de 89,08%;

B) Grupo 3 de documentos:

. no nível 2 de subordinação: média de 91,59%; e

. no nível 3 de subordinação: média de 28,18%;

C) Grupo 4 de documentos:

. no nível 2 de subordinação: média de 82,77%;

. no nível 3 de subordinação: média de 68,99%; e

. no nível 4 de subordinação: média de 42,73%;

D) Grupo 5 de documentos:

. no nível 2 de subordinação: média de 93,84%;

. no nível 3 de subordinação: média de 81,14%;

. no nível 4 de subordinação: média de 38,63%; e

. no nível 5 de subordinação: média de 8,30%.

Através dos cálculos percentuais acima indicados relativos aos Grupos de documentos considerados para esta avaliação, foi verificado que:

- o nível 2 de subordinação dos sumários foi utilizado para a descrição de, no mínimo, 82,77% do conteúdo informativo dos livros e monografias;

- o nível 3 de subordinação dos sumários foi utilizado para a descrição de, no mínimo, 68,99% do conteúdo informativo dos livros e monografias; e

- os níveis 4 e 5 de subordinação dos sumários foram utilizados, respectivamente, para a descrição de, no mínimo, 38,63% e de 8,30% do conteúdo informativo dos livros e monografias.

Com base nos percentuais acima ressaltados, foi possível inferir que os autores dos documentos integrantes da amostra investigada, ao definirem a estrutura e a dimensão discursiva de seus trabalhos, concentraram e relacionaram a abordagem mais específica dos temas centrais nas páginas correspondentes aos níveis 2 e 3 de subordinação dos sumários. Observa-se que, a partir dos níveis 4 e 5 de subordinação, o percentual de páginas de texto correspondente a estes níveis decaiu sensivelmente.

Conseqüentemente, foi possível deduzir que, até o nível 3 de subordinação dos sumários localiza-se a maior concentração de conteúdo informativo de livros e monografias. Os níveis 4 e 5 de subordinação dos sumários, além de ocorrerem de forma mais dispersiva e esporádica nos textos, compreendem um número de páginas bem reduzido e são utilizados para a descrição de tópicos mais específicos do que os descritos até o nível 3 de subordinação.

Complementando as observações relativas aos dados incluídos no Anexo 33, cabe também mencionar as médias de incidências dos níveis 2, 3, 4 e 5 de subordinação nos sumários dos documentos que integraram os Grupos examinados. Estas médias são a seguir discriminadas:

A) Grupo 2 de documentos:

no nível 2 de subordinação: média de 22,5 incidências;

B) Grupo 3 de documentos:

no nível 2 de subordinação: média de 15,8 incidências;

no nível 3 de subordinação: média de 11,3 incidências;

C) Grupo 4 de documentos:

no nível 2 de subordinação: média de 17,0 incidências;

no nível 3 de subordinação: média de 27,0 incidências;

no nível 4 de subordinação: média de 21,0 incidências;

D) Grupo 5 de documentos:

no nível 2 de subordinação: média de 13,5 incidências;

no nível 3 de subordinação: média de 39,5 incidências;

no nível 4 de subordinação: média de 35,5 incidências;

no nível 5 de subordinação: média de 8,0 incidências.

A partir das médias acima apresentadas, foi possível conferir que:

- Nos documentos integrantes do Grupo 2, as incidências do nível 2 de subordinação foram bem expressivas em relação à especificidade da abordagem dos temas centrais dos livros e monografias demonstrada pelos seus sumários (média de 22,5 por documento);

- Nos documentos integrantes do Grupo 3, as incidências do nível 2 de subordinação (15,8) foram superiores às do nível 3 (11,3); e
- Nos documentos integrantes dos Grupos 4 e 5, as incidências do nível 3 de subordinação (27,0 e 39,5, respectivamente) foram superiores às dos demais níveis.

A principal avaliação resultante das médias de incidências de níveis de subordinação nos sumários dos livros e monografias examinados diz respeito à caracterizada tendência dos autores, conforme a especificidade de seus trabalhos, de hierarquizarem a sua estrutura textual, com maior intensidade, até o equivalente ao nível 3 de subordinação. Observa-se para os Grupos 4 e 5 de documentos, que as médias de incidências nos níveis 4 e 5 de subordinação decaíram em relação ao nível 3 de subordinação. Este dado também pode representar uma maior concentração do conteúdo informativo de livros e monografias até o equivalente ao nível 3 de subordinação para estes Grupos de documentos.

Estas duas análises (número de páginas de texto/nível de subordinação, e incidências de níveis de subordinação) vieram ao encontro dos resultados apontados tanto para a pós quanto para a pré-coordenação, quando, como já mencionado, foi verificado que a maior formação de cabeçalhos de assunto ocorreu até os níveis 2 e 3 de subordinação, sendo que especialmente o nível 3 foi mais rico para a identificação de palavras semanticamente representativas em livros e monografias com abordagem mais específica de seus temas centrais (documentos dos Grupos 4 e 5).

Portanto, pelo estudo efetuado, ficou comprovado que existe uma relação proporcional entre o número de cabeçalhos formados nos níveis de subordinação dos sumários de livros e monografias e o número de páginas de texto correspondentes a estes níveis.

Esta relação pode configurar a propriedade de se atribuir cabeçalhos de assunto formados a partir de palavras e/ou expressões disponíveis nos níveis 2 ou 3 de subordinação dos sumários, com elevado grau de representatividade e de pertinência do conteúdo informativo de livros e monografias.

Foi possível deduzir que cabeçalhos de assunto formados conjuntamente a partir dos títulos, subtítulos e sumários são virtual, potencial e adequadamente representativos do conteúdo informativo de livros e monografias, considerando-se, também, a vantagem de serem originados da própria terminologia utilizada por seus autores. Além disso, os títulos, subtítulos e sumários são elementos estruturais que conferem ao profissional da informação opções de identificação de palavras representativas de conteúdo de acordo com a especificidade da indexação a executar, em proporção ao número de páginas em que determinado assunto é apresentado pelo documento.

A identificação desta relação, aliada aos resultados apontados tanto para Sistemas de Informação pós-coordenados quanto para Sistemas pré-coordenados, veio enriquecer e fundamentar a aplicação da técnica descrita e sugerida pela presente Dissertação, a qual objetivou, principalmente, promover o aumento da disponibilidade de recuperação de informações

transmitidas por livros e monografias, documentos de inquestionável importância no âmbito da literatura de cuhos geral e especializado.



CITAÇÕES E NOTAS

- 01 - SOERGEL, Dagobert. Updating and maintenance of indexing languages and *Thesauri*. In: \_\_\_\_\_. *Indexing languages & Thesauri: construction & maintenance*. Los Angeles: Melville Pub. Co., 1974. p. 457-467.
- 02 - WORMELL, Irene. SAP - a new way to produce subject descriptions of books. *Journal of Information Science*, n. 3, p. 39 - 43, 1981.
- 03 - Cf. SOERGEL, *op. cit.*, p. 457.
- 04 - A Rede BIBLIODATA (CALCO) intenciona promover o desenvolvimento dos recursos de recuperação de informações na sua Base de Dados bibliográficos. Esses recursos referem-se, principalmente, à recuperação por assunto. Segundo informações obtidas junto à administração da Rede, seria utilizado um "software" da IBM denominado "STAIRS" o qual permitiria a "varredura" de todos os campos preenchidos pela Bibliotecas e Centros de Documentação cooperantes nas planilhas ou "diskettes" do formato CALCO. A recuperação poderia ser efetuada através da interseção de termos com a aplicação concomitante de parâmetros de filtragem. Esses termos seriam pesquisados e identificados pelo computador em toda a ficha catalográfica do documento. Especificamente quanto à busca nos campos referentes aos cabeçalhos de assunto, haveria a possibilidade de pesquisa

através dos cabeçalhos simples ou compostos, como também através dos cabeçalhos tópicos e dos subcabeçalhos, o que já ocorre na Base de Dados. Outra vantagem seria a possibilidade de elaboração de pesquisas através de "Termos Livres", um mecanismo concebido pelo formato CALCO que permite a indexação por assunto através de termos não incluídos no vocabulário controlado adotado pela Rede mas que seriam apropriados para determinadas Bibliotecas ou Centros de Documentação cooperantes.

05 - Na Biblioteca do BNDES não seria possível efetuar o teste desejado em função das prioridades estabelecidas pela sua Administração relativas à ocupação do tempo de uso do sistema automatizado. No caso da Rede BIBLIODATA (CALCO), como também da sua Base de Dados bibliográficos, a inerente diversidade de características e objetivos das Bibliotecas e Centros de Documentação integrantes da Rede, não permitiria a obtenção de resultados e de avaliações correspondentes a um universo homogêneo.

## 9 -- CONCLUSÕES

A importante e tradicional participação que os livros e, mais recentemente, as monografias possuem no processo de transmissão do conhecimento, as restrições de acesso ao conteúdo informativo destes documentos ocasionadas pelos procedimentos técnicos em geral verificados nos Sistemas de Informação, e, como pesquisado no universo literário da Ciência da Informação, o reduzido número de pesquisas devotadas ao aumento da acessibilidade de seus conteúdos temáticos, foram circunstâncias e problemas enfaticamente abordados nas seções introdutórias da presente Dissertação.

Diante do quadro apresentado, os objetivos definidos para a execução da presente pesquisa foram plenamente atingidos, como a seguir discriminado.

Os títulos, subtítulos e sumários foram elementos estruturais de livros e monografias que, conjuntamente, permitiram a identificação de um significativo número de palavras e/ou expressões com elevada representatividade dos conceitos veiculados pelos documentos examinados, com a vantagem de serem redigidas com o emprego da terminologia corrente do campo Economia.

O método de interpretação conceitual empregado favoreceu, sobremaneira, a identificação de um grande número de palavras, expressões e nomes próprios potencialmente representativos da sobrecidade dos documentos integrantes da amostra, através de seus títulos, subtítulos e sumários, uma vez que o esquema que

orientou todo este processo (ver página 67) foi elucidativo, conferindo facilidade e propriedade para a seleção das palavras.

Cabe mencionar que, tanto a utilização concomitante dos títulos, subtítulos e sumários quanto o método de interpretação conceitual empregado foram recursos capazes de reduzir o esforço intelectual inerente à análise de conteúdo informativo de documentos.

A definição de níveis de subordinação nos sumários dos livros e monografias examinados foi metodologicamente muito oportuna, porque permitiu delimitar com clareza os graus de especificidade de abordagem dos assuntos que os autores conferiram a seus trabalhos, o que, além de facilitar a seleção de palavras, concorreu para que a análise dos dados coletados e dos resultados obtidos fossem expressadas e discutidas em função de dois postulados que, dentre outros, regem a indexação e a recuperação de informações: revocação e precisão. Também os níveis de subordinação estipulados para os sumários orientaram o agrupamento distinto dos 30 documentos integrantes da amostra, conforme a similariedade estrutural e hierárquica de seus sumários, ocasionando uma análise uniforme para livros e monografias com características de distribuição textual semelhantes.

O expressivo número de palavras extraídas conjuntamente dos títulos, subtítulos e sumários durante a análise conceitual para a formação de cabeçalhos de assunto, demonstrou o potencial desses elementos estruturais como fontes apropriadas para a identificação e seleção de palavras com elevada

representatividade conceitual.

Os resultados da pesquisa junto a um Sistema de Informação pós-coordenado, no caso a Biblioteca do BNDES, foram muito significativos, tanto no que se refere aos acréscimos em relação à catalogação por assunto efetuada por aquele Sistema, quanto à demonstração do potencial dos títulos, subtítulos e sumários para a constituição de cabeçalhos de assunto conceitualmente representativos em Sistemas pós-coordenados.

Ficou demonstrado que, para Sistemas que adotem a pós-coordenação na indexação e na recuperação de informações, recorrer até o nível 2 de subordinação dos sumários para livros e monografias com menor grau de especificidade de abordagem de assuntos, e recorrer até o nível 3 de subordinação dos sumários para livros e monografias com maior grau de especificidade, permite a formação de um considerável número de cabeçalhos de assunto pertinentes, os quais, se colocados à disposição dos usuários dos Sistemas, ocasionariam uma maior acessibilidade ao conteúdo informativo real desses documentos, assim como provocariam uma dinamização e uma maximização de seu uso adequado.

Como consequência da pesquisa realizada junto à Biblioteca do BNDES, e com o intuito de fornecer subsídios para o aprimoramento dos serviços prestados por aquele Sistema, cabe sugerir que o emprego da cronologia seja estendido para a indexação e recuperação de informações por assunto de todos os itens bibliográficos a processar, e não somente para tabelas estatísticas, como tem sido efetuado. Já que o Sistema dispõe de

bons recursos operacionais, como a interseção automática de termos de indexação para recuperações por assunto, a disponibilidade da cronologia conferiria maiores opções para a formulação de estratégias de busca, sobretudo em pesquisas com enfoque retrospectivo.

Os resultados da investigação relativos a um Sistema de Informação pré-coordenado, no caso a Rede BIBLIODATA (CALCO), foram também muito expressivos, tanto no que tange os acréscimos em relação à catalogação por assunto executada por aquele Sistema, quanto à demonstração do potencial dos títulos, subtítulos e sumários para a formação de cabeçalhos de assunto conceitualmente representativos.

Também ficou demonstrado que, para Sistemas que adotem a pré-coordenação na indexação e na recuperação de informações, recorrer até o nível 2 de subordinação dos sumários para livros e monografias com menor grau de especificidade de abordagem de assuntos, e recorrer até o nível 3 de subordinação para livros e monografias com maior grau de especificidade, permite a formação de um grande número de cabeçalhos de assunto pertinentes, os quais, se colocados à disposição dos usuários dos Sistemas, ocasionariam os mesmos benefícios já descritos para os Sistemas pós-coordenados.

Reiterando, os resultados obtidos junto aos dois Sistemas de Informação examinados foram surpreendentemente homogêneos, como já mencionado no item 8.3, o que respalda e confirma a aplicabilidade da metodologia empregada na presente pesquisa para se identificar palavras potencialmente representativas do

conteúdo informativo de livros e monografias, visando indexá-los com maior exaustividade e, em paralelo, não comprometer a precisão a alcançar em futuras recuperações.

Outro ponto que também contribuiu para corroborar a aplicabilidade da metodologia empregada, assim como as afirmações efetuadas em consequência dos resultados auferidos junto à pós e à pré-coordenação, foi a análise da quantificação das páginas de texto compreendidas pelos níveis de subordinação delimitados para os sumários dos documentos examinados. Foi constatada a concentração de textos até o correspondente ao nível 2 de subordinação para livros e monografias com menor grau de especificidade de abordagem, e até o nível 3 de subordinação para livros e monografias com maior grau de especificidade de abordagem, níveis estes onde, durante a pesquisa realizada, a formação de cabeçalhos de assunto foi mais intensa para documentos com características estruturais semelhantes.

Este conjunto de resultados e respectivas avaliações permitiu apresentar e sugerir à Comunidade da Informação um método racional, alternativo, de fácil assimilação e execução técnica, capaz de reduzir ou de, quiçá, solucionar as restrições de acesso ao conteúdo informativo de livros e monografias ocorridas nos Sistemas de Informação, independente de sua capacidade operacional e das Políticas de Indexação que adotam. A versatilidade do método é conferida pelas opções de se indexar com menor ou maior especificidade, ou com menor ou maior exaustividade, de acordo com os níveis de subordinação estabelecidos em função da estrutura textual dos livros e monografias demonstrada por seus sumários.

Como enfaticamente abordado, as limitações de acesso ao conteúdo informativo de livros e monografias, em geral verificadas nos Sistemas de Informação, puderam ser retratadas e confirmadas através das mensurações dos cabeçalhos de assunto atribuídos tanto pela Biblioteca do BMDIS quanto pela Rede BIBLIODATA (CALCO) para os documentos integrantes da amostra. Esses Sistemas deveriam promover a revisão de suas Políticas de Indexação, independente de adotarem ou não as soluções propostas pela presente Dissertação, uma vez que, além de restringirem a seus usuários o acesso ao conteúdo informativo total e real dessas categorias de material bibliográfico, o retorno dos investimentos efetuados para a sua aquisição, processamento técnico e manutenção certamente atingiu níveis aquém do desejáveis.

O aumento da disponibilidade de acesso ao conteúdo informativo de livros e monografias demonstrado pelos resultados obtidos junto aos dois Sistemas de Informação examinados, certamente possibilitaria a maximização e a otimização do uso destas coleções, item indispensável para o cumprimento de seus objetivos fim. Estende-se esta constatação a todos os demais Sistemas de Informação instituídos no País, pois, como já mencionado, é preciso tornar a informação textual mais acessível para todas as camadas sociais, independente do propósito de sua utilização, como um dos componentes da evolução social, científica, tecnológica e cultural de nossa população e, principalmente, para colaborar no processo educacional de níveis elementar, intermediário e superior.



Disvirtuando-se do principal enfoque da presente Dissertação, não poderiam ser omitidas algumas observações extraídas durante a compatibilização dos cabeçalhos de assunto formados a partir da análise conceitual com os vocabulários controlados adotados pelos Sistemas de Informação investigados.

As linguagens de indexação utilizadas pela Biblioteca do BNDES e pela Rede BIBLIODATA (CALCO) possuíam uma capacidade de representação de assuntos inferior ao número de conceitos relevantes tratados pelos documentos integrantes da amostra examinada. Na Biblioteca do BNDES, mesmo com os recursos de interseção automática de termos de indexação, o vocabulário controlado não atendeu, de modo ideal, às necessidades de representação conceitual levantadas durante a execução da pesquisa (42,11% de defasagem). Na Rede BIBLIODATA (CALCO) também foi apontada uma considerável defasagem do vocabulário controlado (34,86%).

Estes indicadores, além de comprometerem a qualidade do processamento técnico dos documentos e do atendimento aos usuários dos Sistemas investigados, permitem sugerir uma revisão e uma atualização dessas linguagens de indexação, principalmente porque ficou caracterizada uma reduzida proximidade com a linguagem natural empregada pelos autores dos documentos examinados.

Na Biblioteca do BNDES seria fundamental efetuar uma revisão e uma atualização do vocabulário controlado, visto que o elevado índice de defasagem apontado (42,11%) referiu-se ao

campo Economia, justamente aquele de maior importância para o contexto da instituição à qual atende.

Para a Rede BIBLIODATA (CALCO), uma revisão e atualização do vocabulário controlado ocorreria bem a propósito, uma vez que aquele Sistema vem promovendo freqüentes aperfeiçoamentos voltados para a recuperação de informações por assunto (cf. Nota 04 do item 8.2.3.2.2). Com o futuro emprego da interseção de termos para pesquisas na Base de Dados bibliográficos, pode-se inferir uma virtual tendência da Rede em desvencilhar-se do uso de um vocabulário controlado oriundo do LCSH - Library of Congress Subject Headings, cujas regras e padrões para linguagens de indexação pré-coordenadas, impõem uma série de restrições que limitam a busca e o acesso a informações precisas.

Cabe mencionar que a utilização freqüente de palavras extraídas de elementos estruturais de livros e monografias (como os títulos, subtítulos e sumários), ou mesmo de outras categorias de documentos, para o processo de indexação também seria um mecanismo capaz de identificar a terminologia corrente e conceitos recentemente instituídos em determinada literatura para posterior incorporação criteriosa em vocabulários controlados.

Todos os resultados e avaliações apresentados oferecem subsídios suficientes para que, em pesquisas posteriores, seja efetuada uma análise de custo-benefício da aplicação da metodologia empregada na presente Dissertação, uma vez que ela naturalmente implicaria em elevação de dispêndios operacionais dos Sistemas de Informação.

Independente do fator custo, é razoável afirmar que o aprimoramento dos serviços prestados pelos Sistemas de Informação, ocasionado pela maior acessibilidade do conteúdo informativo real de documentos tão valiosos como os livros e as monografias, traria significativos benefícios, tanto para a sociedade a que se propõem a atender (na medida em que tornar-se-iam mais próximos de diferentes necessidades de informação e de diferentes categorias de usuários), quanto para os próprios Sistemas, estes especificamente no sentido de se afirmarem como elementos essenciais do processo de transmissão do conhecimento.

Reportando-se à futura capacitação operacional das Bibliotecas e Centros de Documentação, dentre as novas tecnologias da informação disponíveis para o aprimoramento de serviços encontram-se os "scanners", equipamentos para leitura de textos e posterior gravação em discos óticos. A leitura de títulos, subtítulos e sumários de livros e monografias através de "scanners", seguindo-se a metodologia empregada na presente pesquisa, com gravação em discos óticos destinados a um Sistema programado para indexação automática, facilitaria não somente a operacionalização do processamento técnico, mas também permitiria um acesso mais imediato e específico às informações contidas nestes documentos.

Esta Dissertação intencionou apresentar e sugerir soluções para um problema apontado, o que foi plenamente alcançado, além de ser uma contribuição inédita para a Ciência da Informação a nível nacional e, porque não dizer, também a nível

internacional, uma vez que a revisão da literatura indicou apenas a existência de pesquisas semelhantes. Entretanto, sua intenção maior foi deixar registrada a necessidade dos Sistemas de Informação alterarem sua conduta relativa ao processamento técnico de livros e monografias, uma vez que, em países como o Brasil, o acesso adequado e preciso à informação é um privilégio para uma minoria de suas populações, e onde os livros ainda são documentos culturalmente instituídos como aqueles de maior aceitação, confiabilidade e popularidade. A forma física do livro representa, para muitas pessoas, estudo, aprendizado, crescimento individual e coletivo.

A informação tornou-se um insumo também econômico, e, na atual conjuntura, não é aconselhável abstrair-se desta realidade. Portanto, quanto maior for o acesso e a assimilação profícua do conhecimento, maior será o preparo da sociedade para responder positivamente às freqüentes transformações a que é submetida, nem sempre favoráveis.

10 REFERÊNCIAS BIBLIOGRÁFICAS

ATHERTON, Pauline. Improved subject access to books in on-line catalogs. In: INTERNATIONAL ON-LINE INFORMATION MEETING, 1st., London, 1977. [Proceedings] Oxford: Learned Information, 1977. 234 p. p. 131-138.

ATHERTON, Pauline et al. *Books are for use: final report to the Council on Library Resources*. Syracuse, NY: Syracuse University, School of Information Studies, 1970. 172 p.

BAILLEY, Kenneth D. *Methods of social research*. New York: Free Press, 1978. p. 84.

DARUGA, Denis Borges. O avanço do feudalismo informacional. *Jornal do Brasil*, Rio de Janeiro, 14 jul. 1991. Caderno Idéias, p. 10-11.

DELL, Daniel. *O advento da sociedade pós-industrial: uma tentativa de previsão social*. São Paulo: Cultrix, 1977. p. 139.

DYRNE, Alex, MICCO, Mary. Improving OPAC subject access: the ADFR experiment. *College and Research Libraries*, v.49, n.5, p. 432-441, Sep. 1988.

COCHRANE, Pauline Atherton. The management of subject cataloguing in an era of interactive catalogues and bibliographic networks. *Cataloguing Australia*, v.11, n.4, p. 2-8, Dec. 1985.

..... Modern subject access in the online age. *American Libraries*, v.15, p. 336-339, May 1984.

..... *Redesign of catalogs and indexes for improved online subject access: selected papers of Pauline A. Cochrane*. Phoenix: Dryx Press, 1985. 484 p.

CHRISTOVÁO, H.T. The aging of the literature of biomedical sciences in developed and developing countries. *Scientometrics*, v.7, p. 411-430, 1985.

CLEVELAND, Donald B., CLEVELAND, Ana D. Indexing methods and procedures. In: ..... *Introduction to indexing and abstracting*. Littleton, Colo.: Libraries Unlimited, 1983. p. 62-75.

DEHART, Florence E., REITSMA, Richard. Subject searching and table of contents in single-work titles. *Technical Services Quarterly*, v.7, n.1, p. 33-51, 1989.

DIODATO, Virgil. Table of contents and book index: how well do they match readers' descriptions of books? *Libraries Resources & Technical Services*, v.30, n.4, p. 402-412, Oct./Dec. 1986.

ETCHINSKI, George. Enhancing subject data: talk or action. *Cataloguing Australia*, v.11, n.4, p. 9-18, Dec. 1985.

ENCYCLOPAEDIA Britannica. Chicago, Ill. : Britannica Inc., 1964. 23 v. v.3: Dalfour to Both.

FUNDAÇÃO GETÚLIO VARGAS. Biblioteca Central, FUNDAÇÃO GETÚLIO VARGAS. Centro de Processamento de Dados. BIBLIODATA (CALCO). *Cabecalhos de assunto: normas adotadas pela Rede BIBLIODATA/CALCO*. Rio de Janeiro: 1989. 1 v.

FUNDAÇÃO GETÚLIO VARGAS. Centro de Processamento de Dados. Gerência de Sistemas de Acesso Público. *CALCO on-line: manual do usuário* [Rio de Janeiro: 1989?] 22 f.

GRIFFITH, Delver C., SERVI, Patricia N., ANKER, Anita L. et al. The aging of scientific literature: a citation analysis. *Journal of Documentation*, v.35, n.3, p. 179-196, Sep. 1979.

INGWERSEN, Peter, WORMELL, Irene. Means to improve subject access and representation in modern information retrieval. *Libri*, v.38, n.2, p. 94-119, June 1988.

LANCASTER, F. W., ELLIKER, Calvin, CONNELL, Tschera Harkness. Subject analysis. *Annual Review Of Information Science & Technology*, v.24, p. 35-84, 1989.

LANDRIDGE, Derek. *Approach to classification: for students of librarianship*. London: Clive Bingley, 1973. p.61.

LUCAS, Clarinda Rodrigues. O papel do Sistema de Informação no processo de transferência de tecnologia. *Ciência da Informação*, v.16, n.2, p. 181-191, jul./dez. 1987.

MANDEL, Carol A. Enriching the library catalog record for subject access. *Library Resources & Technical Services*, v.29, n.1, p. 5-15, Jan./Mar. 1985.

MANDEL, Carol A., HERSCHMAN, Judith. Online subject access: enhancing the library catalog. *Journal of Academic Librarianship*, v.9, p. 140-155, July 1983.

MATSUI, Sachiko, ASANO, Chooichiro, MATSUUDA, Yoshiro. A system for generating publication statistics based on bibliographic information: bibliometric analysis for the development of economic and social thought. *Hitatsubashi Journal of Economics*, v.30, n.2, p. 121-156, Dec. 1989.

MICCO, Mary. *An exploratory study of three subject access systems in medicine: LCSH, MeSH, PRECIS*. Pittsburgh: University of Pittsburgh, 1980. (Doc. Thesis)

MICCO, Mary, RAMBLER, Linda. *Cognitive overload: an emerging problem hypertext for improved subject access: a proposed solution* [s.l.: s.n.] 1989. 8 p.



- MICCO, Mary, SMITH, Irma, HSIAO, Su-Ann et al. Knowledge representation: subject analyses. *Library Software Review*, p. 82-87, Apr. 1987.
- MILSTAED, Jessica L. *Subject access systems: alternatives in design*. New York: Academic Press, 1984. 212 p.
- NEILL, S.D. The dilemma of subjective in information organization and retrieval. *Journal of Documentation*, v.43, n.3, p. 193-211, Sep. 1987.
- POULSEN, Claus. Subject access to new subjects, specific paradigms and surveys: PARADOKS-registration. *Libri*, v.40, n.3, p. 179-202, 1990.
- PRICE, Derek J. de Solla. *O desenvolvimento da ciência: análise histórica, filosófica, sociológica e econômica*. Rio de Janeiro: Livros Técnicos e Científicos, 1976. 77 p.
- RADA, Juan. Information technology and the Third World. In: FORESTER, Tom, ed. *The information technology revolution*. Cambridge, Mass.: MIT Press, 1985. p. 571-589.
- RANGANATHAN, S. R. *Colon classification*. Bombay: Asia Publishing House, 1964. p. 1.25.

RAWSKI, Conrad H. The scientific study of subject literatures. *Visible Language*, v.12, p. 5-23, Winter 1977.

RUSSEL, Keith W., comp. *Subject access*: report of a meeting sponsored by the Council on Library Resources, 1982, June 7-9, Dublin, OH. Washington: Council on Library Resources, Inc., 1982. 108 p.

SARACEVIC, Tefko. *Information science*: origin, evolution and relations. Pre-print, 1991. 16 p.

SCHWARTZ, Candy, EISENMANN, Laura Malin. Subject analysis. *Annual Review of Information Science & Technology*, v.21, p. 37-61, 1986.

SETTEL, Barbara, ed. *Subject description of books*: a manual of procedures for augmenting subject descriptions in library catalogs. Syracuse, NY: Syracuse University, School of Information Studies, 1977.

SETTEL, Barbara, COCHRANE, Pauline A. Augmenting subject description for books in online catalogs. *Database*, v.5, n.4, p. 29-37, Dec. 1982.

SERGEL, Dagobert. Updating and maintenance of indexing languages and *Thesauri*. In: \_\_\_\_\_ *Indexing languages & Thesauri*: construction & maintenance. Los Angeles: Melville Pub. Co., 1974. p. 457-467.

TELL, Bjorn V. Deep analyses of book contents: fact retrieval from Government Committee reports using the SAP-technique online. In: ESSEN SYMPOSIUM, Essen, 1981. [Proceedings] Essen: Gesamthochschulbibliothek, 1982. 5.v. v.3: Future of library networks, p. 116-126.

TELL, Bjorn V., WORMELL, Irene. Factual data retrieval from a data base constructed according to the SAP-technique: a project incorporating Government Committee reports. In: INTERNATIONAL RESEARCH FORUM IN INFORMATION SCIENCE, 4., Borås, Sweden, 1981. *Proceedings*. Lund: Lund University, 1981. p.135-149.

TRAVIS, Irene L., FIDEL, Raya. Subject analysis. *Annual Review of Information Science & Technology*, v.17, p. 143-157, 1982.

VICKERY, B.C. *La classification à facettes: guide pour la construction et l'utilisation de schémas spéciaux*. Paris: Gauthier-Villars, 1963. 60 p.

\_\_\_\_\_. *Classification and indexing in science*. 2.ed. London: Butterworths Scientific Publications, 1959. 235 p.

\_\_\_\_\_. The description of documents. In: \_\_\_\_\_. *On retrieval system theory*. 2nd. éd. London: Butterworths, 1968. p.20-32.

VICKERY, B.C.. Knowledge representation: a brief review. *Journal of Documentation*, v.42, n.3, p. 145-159, Sep. 1986.

----- The processes of subject indexing. In: *Techniques of information retrieval*. London: Butterworths, 1970. p.76-86.

WORMELL, Irene. SAP: a new way to produce subject description of books. *Journal of Information Science*, n.3, p. 39-43, 1981.

----- SAP: a project for subject description of books. *Tidsk. Dokum.*, v.34, n.1, p. 3-6, 1978.

----- *Subject Access Project-SAP: improved subject retrieval for monographic publications*. Lund: Lund University, 1985.  
(Doc. Thesis)