

# O CABEÇALHO DE ASSUNTO DA REDE BIBLIODATA/CALCO: USO E RECUPERAÇÃO NA BASE ACERVUS/UNICAMP

Nirlei Maria OLIVEIRA  
UNICAMP

Maria das Dores Rosa ALVES  
UNICAMP

Gilmar VICENTE  
UNICAMP

## RESUMO

O catálogo on-line é o maior impacto que se deu junto ao usuário com a possibilidade de busca rápida de informação. Analisando a base de dados Acervus, especificamente o cabeçalho de assunto, seu uso e recuperação tanto na perspectiva do usuário quanto do bibliotecário, foram utilizados 97 registros, sendo apresentadas as respectivas obras ao juiz 1, bibliotecário e a juizes docentes para a representação do conteúdo. Os resultados apresentam relevantes aspectos dos índices de concordância, trazendo à reflexão quanto às formas de uso do cabeçalho de assunto na recuperação da informação e satisfação do usuário.

**Palavras-chave:** Cabeçalho de assunto; Recuperação da informação; Cabeçalho de assunto - utilização.

## INTRODUÇÃO

Produtos e serviços automatizados são constantemente desenvolvidos nas bibliotecas universitárias. O catálogo *on-line* foi

um dos primeiros serviços automatizados e de maior impacto nas bibliotecas.

Este impacto pode ser visualizado em dois segmentos: no processamento técnico e na área de serviços ao público, tanto no que diz respeito ao aprimoramento e atualização de catálogos quanto na realização de buscas por meio de pontos de acesso à informação (Zanaga, 1994).

O maior impacto ocorreu junto ao usuário com a possibilidade de busca rápida de informações. Reforçando esta questão cabe ressaltar o estudo de Alzofon & Polis (1984) que constataram que 95% dos usuários da Ohio State University utilizam o catálogo *on-line* como primeira fonte de informação.

Em outra pesquisa, Cardoso & Gardin (1994) constataram que a busca de informação dos usuários da Biblioteca Central da UNICAMP prioritariamente é o catálogo *on-line*.

É indiscutível o uso do catálogo automatizado, no entanto, há que se considerar questões como satisfação do usuário e padrões de busca de informação. A satisfação do usuário apresenta-se como uma questão complexa, pois é multidimensional.

Thorne & Whitlatch (1994) em uma pesquisa na biblioteca da San José State University analisaram a satisfação do usuário em relação ao catálogo automatizado e 75% dos sujeitos se mostraram totalmente satisfeitos e apenas 5,4% encontravam-se insatisfeitos com o catálogo automatizado. Uma questão interessante nesta pesquisa é que 51,4% contavam com a assistência de um profissional de referência.

Outra questão ligada à satisfação do usuário está intrinsecamente conectada à utilização dos padrões de busca de informação. Em pesquisa de Cherry (1994), Moore (1981), Alzofon & Polis (1984), os usuários apresentam um padrão único de busca de informação: autor, título, item conhecido, assunto. Em todas estas pesquisas, o assunto apresentou-se como uma questão problemática.

Para Alzofon & Polis (1984), a pesquisa por assunto requer o uso correto do cabeçalho de assunto da Library of Congress e neste ponto o catálogo automatizado é falho.

Moore (1981) reforça esta assertiva e sugere a necessidade de oferecer busca utilizando expressão booleana, palavras do título, explicitar a estrutura do assunto no catálogo *on-line*, bem como desenvolver treinamento que capacite o usuário a utilizar com eficiência o catálogo.

Um estudo realizado por Pritchard (1979), na Library of Congress, revelou que 70% dos usuários desejavam acesso ao catálogo por assunto. Ocorre o que Cherry (1994) diz: o usuário tem apenas um conhecimento rudimentar ou desconhece a estrutura do cabeçalho de assunto, daí a utilização de um item conhecido na pesquisa. Uma questão relevante é a resistência do usuário em solicitar o auxílio do bibliotecário sobre o uso do cabeçalho de assunto.

A preocupação com o resultado zero na busca por assunto, isto é assunto não localizado, levou a biblioteca a preocupar-se com a educação dos usuários, programas de conversão de resultados zero em outros tipos de busca. Cherry (1994) sugere duas soluções para a questão do resultado zero na pesquisa por assunto, uma delas seria o próprio sistema de emitir mensagem sugerindo novas tentativas recorrendo a outras formas de acesso.

Outra sugestão seria do sistema converter automaticamente a busca por título, palavras do título e executá-la, emitir mensagem que o assunto não foi recuperado e remeter para a utilização de outros termos.

O acesso à coleção por assunto parece ser limitado como visto nas pesquisas anteriormente citadas. É por certo uma questão provocativa, tendo em vista que em qualquer produto ou serviço o objetivo fim é o usuário.

Pensando na eficiência da recuperação da informação, o Sistema de Bibliotecas da UNICAMP iniciou a participação na rede de catalogação cooperativa em 1989, gerando uma base local, disponível à partir de 1992. Atualmente a Base Acervus disponibiliza 50% do acervo das bibliotecas da UNICAMP ao usuário. Está em projeto a conversão retrospectiva do restante da coleção.

A participação na Rede Bibliodata/Calco parte da aceitação de determinadas normas e procedimento. A aceitação e utiliza-

ção do cabeçalho de assunto desenvolvido pelos integrantes da rede é uma delas.

Assim, a Rede Bibliodata/Calco utiliza uma linguagem pré coordenada, fundamentada na Library of Congress Subject Headings LCSH - respeitando particularidades da Língua Portuguesa. A razão da opção pela LCSH foi devido a sua multidisciplinaridade, por ser uma lista confiável, uma vez que os termos são estudados por especialistas de diversas áreas (Manual....FGV, 1995).

A linguagem de indexação é um ponto crítico, tendo em vista, que esta reflete na possibilidade de recuperação da informação. Neste ponto é que o usuário vai sentir se o sistema reflete ou não os conteúdos, em uma linguagem comum ao seu universo.

O Manual de cabeçalhos de assunto: normas e procedimentos de 1995 traz os termos eleitos pela Rede Bibliodata Calco, dentro de uma sintaxe pré estabelecida, objetivando a padronização, atribuição e recuperação por assunto.

Segundo este mesmo manual, a lista compilada de cabeçalho de assunto conta hoje com 32.500 cabeçalhos tópicos e suas subdivisões, referências, remissivas, termos LC organizados em ordem alfabética.

A utilização deste vocabulário é visualizada pelos usuários na recuperação da informação. O sistema de recuperação da informação bibliográfica do Sistema de Bibliotecas da UNICAMP é composto dos seguintes módulos: pesquisa, exibição, refinamento, ordenação, resumo das consultas realizadas e auxílio.

O que interessa neste artigo é o módulo pesquisa, pois permite ao usuário recuperar a informação por assunto, autor, título e termo livre.

A pesquisa no campo Assunto está relacionada a cabeçalhos de assunto da Rede Bibliodata/Calco, assim lista todos os documentos contendo o termo pesquisado no campo Assunto.

Pode-se ainda usar a raiz da palavra seguida do sinal \$, este lista todos os termos no campo assunto que comece por esta raiz. Já o sinal \$nº 1,2,3, lista palavras com a mesma raiz e o número de caracteres indicados.

No campo de autor, pode-se usar nome e sobrenome em qualquer ordem, omitindo termos como de, da etc.

A pesquisa em campo múltiplo por ex.: assunto e autor, também é viabilizada.

O usuário também pode fazer pesquisa utilizando operadores e/ou recupera os 2 termos indicados.

Há também a pesquisa no campo livre. Este campo é bastante abrangente e busca o termo em qualquer parágrafo.

Esta pesquisa tem por objetivo analisar a Base Acervus, especificamente o cabeçalho de assunto, tanto na perspectiva do usuário quanto do bibliotecário, no que diz respeito a concordância na utilização de vocabulário para representar o conteúdo das obras.

## MÉTODO

### Sujeitos

Dois tipos de juízes foram usados. Juiz Bibliotecário (J1) um profissional da área com 9 anos de experiência em processamento técnico. Juizes Docentes (J2) - 21 docentes com formação nas áreas de conteúdos pesquisados.

### Material

Para a realização deste estudo foram utilizados os registros bibliográficos constantes na base de dados Acervus do Sistema de Bibliotecas UNICAMP.

Do total de dados inseridos na Base Acervus, retirou-se uma amostragem de 0,01% o que resultou em 97 registros. A partir deste dado procedeu-se a um sorteio sistemático das subclasses o qual valeu-se do terceiro sumário do Dewey Decimal Classification (1971, p. 451).

Assim, foram digitados o número correspondente à subclasse e na recuperação utilizou-se impreterivelmente o primeiro

registro que correspondesse à busca em questão. Após a recuperação do registro, foi extraído o número de chamada, CPD (nº de identificação da obra na Rede Bibliodata/CALCO), assunto e a biblioteca, no qual o mesmo se encontrava. De posse destes dados, por meio de correspondência, foram solicitadas às bibliotecas seccionais o envio das obras selecionadas.

## Procedimento

Estas obras foram enviadas ao J1, ao qual foi solicitado que fizesse a representação do conteúdo das obras utilizando o cabeçalho de assunto da Rede Bibliodata/Calco.

Após este procedimento, estas mesmas obras foram analisadas pelos J2, com formação nas áreas correspondentes aos documentos. Enviou-se um formulário no qual se perguntava quais os termos, expressões, palavras que ele usaria para representar o conteúdo daquela obra.

## RESULTADOS E DISCUSSÃO

De posse destes dados, procedeu-se a análise do índice de concordância entre a Base Acervus, J1 (bibliotecário) e J2 (Docentes) no que diz respeito ao cabeçalho de assunto.

Para efeito de análise, operacionalizou-se da seguinte forma: acordo total = 2 pontos (exatamente o mesmo cabeçalho de assunto; acordo parcial = 1 ponto (concordância em apenas um item, subdivisão, assunto, geográfico, cronológico e forma); desacordo = 0 (cabeçalho de assunto totalmente diferente).

O primeiro aspecto analisado foi o índice de concordância entre Base Acervus X J1 (bibliotecário) no que diz respeito ao cabeçalho de assunto. Obteve-se um índice de concordância de 69,34%.

Este resultado pode indicar duas direções, primeiro um razoável nível de concordância na utilização de cabeçalhos de

assunto inter-grupos, ou ainda, que a utilização de vocabulário controlado aumenta a coerência na representação do conteúdo temático. Segundo, considerando-se que o ideal no índice de concordância seria acima de 75%, ou seja, falta um melhor conhecimento do vocabulário, ou ainda a utilização de dicionários, glossários, manuais, tendo em vista que se espera que estes ajudem a melhorar a coerência. Outra questão que pode ser afetada é a utilização de um vocabulário geral, pois quanto mais específico maior a possibilidade de coerência na descrição, além do que tópicos concretos são mais fáceis de serem indexados do que tópicos abstratos.

É claro que formação acadêmica, interesses, conhecimento especializado, treinamento podem influenciar na leitura e descrição que cada profissional faz de uma obra.

A avaliação do índice de concordância por classes apresenta-se da seguinte forma no quadro 1: as classes 800 e 900 ficaram respectivamente com 100% de concordância; a classe 500 - com 84,61%; a classe 200 com 83,33%; e a classe 100 com 81,81%.

**Quadro 1** - Base Acervus X J1 quanto a cabeçalho de assunto

CLASSES	I. CONCORDÂNCIA
000 - Generalidades	66.66
100 - Filosofia	81.81
200 - Religião	83.83*
300 - Ciências Sociais	37.50
400 - Linguagem	73.33
500 - Ciências Puras	84.61*
600 - Ciências Aplicadas	57.14
700 - Artes	33.33
800 - Literatura	100.00*
900 - História	100.00*

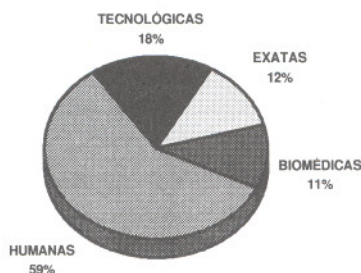
\*acima de 75%.

Nestas cinco classes, com pontuação acima de 75%, pode ter ocorrido o índice significativo dado a especificidade dos conteúdos

e os cabeçalhos permitirem uma representação coerente, ou seja, História, Literatura, Filosofia e Religião são áreas que apresentam quadros bastante estáticos e com subdivisões bem definidas. O mesmo pode ter ocorrido na classe 500 por ser uma área que apresenta conteúdos concretos.

Já a classe 300 com 37.50% e a classe 700 com 33,33% merecem atenção e um estudo detalhado. Na classe 300 por ser área de ciências sociais, há um dinamismo, e multidisciplinaridade intensa, o que pode dar margem a uma gama maior de opção na descrição com conteúdo temático das obras. No caso da classe 700 há uma tabela de classificação própria da área de artes na UNICAMP. Os profissionais alegam que se utilizam desta tabela com números de classificação geral e desta forma o cabeçalho de assunto também fica geral, mesmo tratando-se de uma obra com assunto específico.

Como mostra a figura 1 a área de humanas integra um dos maiores acervos do Sistema de Bibliotecas da UNICAMP, talvez aí haja também uma possível explicação para a dispersão no resultado, além das questões acima citadas.



**Figura 1** - Acervo Sistema de Bibliotecas - UNICAMP

O segundo aspecto analisado refere-se ao índice de concordância entre a Base Acervus e os J2 (docentes). O índice foi de 67%, indicando desvio de comunicação entre a base e seu potencial usuário.

Este dado refere-se a busca de informação e parece refletir as pesquisas realizadas em outras instituições. No estudo de Cherry (1994) a dificuldade com o cabeçalho de assunto da LCSH na



University of Toronto Library é atribuído a desatualização do cabeçalho, ou seja, demora de incorporação de novos termos. Apenas 40% dos usuários disseram que o cabeçalho de assunto era apropriado para representar o conteúdo que pesquisavam.

Na pesquisa de Thorne & Whitlatch (1994, p. 486), 64% dos usuários da Library at San José State University apresentaram dificuldade em utilizar o cabeçalho de assunto, apenas 18%, recupera a informação sem problemas, portanto pior situação que o registrado no presente estudo.

Outros estudos também apontam dificuldades na utilização do cabeçalho de assunto. Assim parece que se utiliza um vocabulário e o usuário fala outro. A disponibilização dos cabeçalhos de assunto poderia evitar muitos resultados zero neste campo. Pode-se dizer que estar disponível um registro na base não configura num acesso fácil ao documento. O usuário tem de percorrer trilhas, até chegar ao caminho desejado, o que não é recomendável para a eficiência do sistema.

No quadro 2, os dados mostram que a classe 900 obteve 100%, enquanto que as classes 300, 500 e 700 obtiveram 80% do índice de concordância, portanto adequados mas nas demais foi insatisfatório.

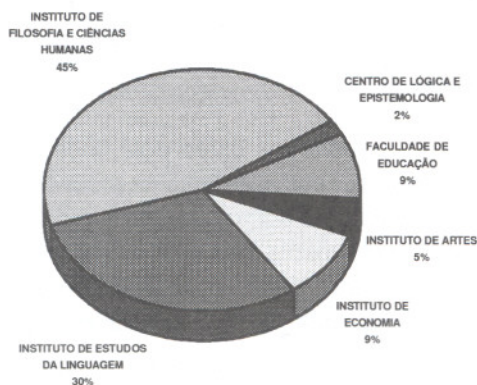
**Quadro 2** - Base Acervus X J2 quanto a cabeçalho de assunto

CLASSES	I. CONCORDÂNCIA
000 - Generalidades	43.00
100 - Filosofia	60.66
200 - Religião	66.66
300 - Ciências Sociais	80.00
400 - Linguagem	50.00
500 - Ciências Puras	80.00*
600 - Ciências Aplicadas	62.50
700 - Artes	80.00*
800 - Literatura	50.00
900 - História	100.00*

\*acima de 75%.

Muitas questões podem ser levantadas. Para Chan (1977), alguns termos são mais populares para determinadas classes de pessoas, muitos termos tornam-se obsoletos, outros adquirem novos significados, ou seja, o vocabulário deveria passar por revisões constantes.

Visualizando as classes dentro do acervo que obtiveram pontuação baixa - 400 e 800 percebe-se que representam apenas 30% do acervo, partindo do pressuposto que sua concentração se dá no Instituto de Estudos Lingüísticos, com comportamento idêntico à classe 300 e estudos anteriormente citados.



**Figura 2** - Acervo área de Humanas

Na análise do índice de concordância de J1 X J2 obteve-se um índice de 58%. Este dado pode ter a mesma análise anterior. As classes 900 e 500 com 90% cada uma; a classe 200 com 88,90%; e a classe 400 e 600 com 75,00% como mostra o quadro 3, apresentam-se com índice ótimo de comunicação entre os profissionais nestas áreas, o bibliotecário representa o conteúdo e o usuário utiliza o mesmo vocabulário para recuperá-lo.

**Quadro 3** - J1 X J2 quanto a cabeçalho de assunto

CLASSES	I. CONCORDÂNCIA
000 - Generalidades	71.00
100 - Filosofia	50.00
200 - Religião	88.90*
300 - Ciências Sociais	50.00
400 - Linguagem	75.00*
500 - Ciências Puras	90.00*
600 - Ciências Aplicadas	75.00*
700 - Artes	60.00
800 - Literatura	40.00
900 - História	90.00*

\*acima de 75%.

Nas classes restantes estes dois tipos de profissionais ainda carecem de um maior entrosamento no processo de comunicação. O uso de um vocabulário comum facilita a rapidez na busca da informação. Moore (1981) aponta uma média de 2 a 8 minutos no uso da máquina para pesquisa. Assim, um bom cabeçalho de assunto economiza tempo da máquina, do bibliotecário e do pesquisador.

Analisou-se também o índice de concordância entre a Base Acervus e o Juiz 1 - Bibliotecário no que diz respeito à classificação, que resultou em um índice de concordância de 92,10%. Apesar de não ser o objeto de estudo, foi considerado oportuno agregar esses dados ao presente trabalho.

Este índice elevado demonstra um alto grau de coerência, o qual talvez seja resultado do conhecimento aprofundado e o bom manuseio do instrumento de classificação CDD, bem como muito tempo desenvolvendo esta atividade leva a uma coerência entre ambos.

A avaliação do índice de concordância dentro de cada classe apresenta-se da seguinte forma no quadro 4.

**Quadro 4 - Número de classificação Índice de concordância / Base Acervus X Juiz 1**

CLASSES	I. CONCORDÂNCIA
000 - Generalidades	71.42
100 - Filosofia	50.00
200 - Religião	80.95*
300 - Ciências Sociais	70.00
400 - Linguagem	50.00
500 - Ciências Puras	78.94*
600 - Ciências Aplicadas	89.47*
700 - Artes	80.00*
800 - Literatura	100.00
900 - História	66.66

\*acima de 75%.

As classes que obtiveram maior índice de concordância foram: 800 - com 100%; 600 - com 84.47%; 200 - com 80.95%; 700 - com 80.00%; e a classe 500 com 78.94%.

As classes restantes ficaram com índice abaixo de 75%, são as seguintes: 000 - com 71.42%; classe 100 e 400 - com 50.00% respectivamente; classe 300 - com 70.00%; classe 900 - com 66.66%. Elas precisariam de uma revisão já que nem mesmo os profissionais da área conseguem o índice desejado.

Este resultado pode ter ocorrido por diversos fatores: por estas serem áreas do conhecimento em que a multidisciplinariedade atinge com mais rapidez e os instrumentos de classificação demoram para inserir novos assuntos. Ou ainda, os assuntos podem ser classificados com flexibilidade dentro de cada classe.

Além destes dados, verificou-se, ainda, que dos 97 cabeçalhos de assuntos extraídos da Base Acervus apenas 17 não eram utilizados ou foram utilizados incorretamente, sendo que nove foram implantados pela Biblioteca da UNICAMP, possivelmente para atender a demanda específica.

## CONCLUSÃO

Este trabalho tentou fornecer dados preliminares sobre o cabeçalho de assunto e fornecer algumas respostas para questões iniciais.

A análise geral dos dados aponta um bom uso do cabeçalho de assunto da Rede Bibliodata/Calco, bem como um bom índice de recuperação pelos usuários em algumas classes do conhecimento. Por outro lado, apresenta áreas em que a comunicação não se encontra em um nível desejável.

Outro ponto que se pode concluir é que há uma coesão inter-grupos, que reflete também na recuperação da informação e satisfação do usuário. No entanto a área de Humanas necessita de estudos para melhorar a descrição de conteúdos e efetivamente facilitar a recuperação da informação. Conclui-se, ainda, que o vocabulário utilizado permite uma aproximação bastante eficiente da linguagem do pesquisador/usuário nas áreas Exatas, Tecnológicas e Biomédicas.

Recomenda-se formar grupos de estudo para aprofundar o tema; disponibilizar o cabeçalho de assunto para o usuário, o que poderá aumentar o uso deste instrumento para eficiência na recuperação da informação; integrar os setores de processamento técnico e referência, para verificação do uso da linguagem de indexação e incorporar a lista de cabeçalho de assunto da Rede Bibliodata/CALCO de vocabulários especializados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALZOFON, S. R. & VAN POLIS, N. Patterns of searching and success rates in an on-line public access catalog. **College & Research Libraries**, March, p.110-115, 1984.
- CARDOSO, I. V. & GARDIN, E. de L. **Setor de referência**: comportamento de usuários da Biblioteca Central da UNICAMP. Campinas: [s.n.], 1994. (no prelo).

- CHAN, L. M. The principle of uniform heading in Library of Congress Subject Headings. **Library Resources & Technical Services**, v.22, n.2 spring, p.127-135, 1978.
- CHERRY, J. M. Improving subject access in OPACS: an exploratory study of conversion of user's queries. **The Journal of Academic Librarianship**, v.18, n.2, p.95-99, 1994.
- DEWEY, M. **Dewey decimal classification and relative index**. New York: Forest, 1971.
- MANUAL DE CABEÇALHOS DE ASSUNTO**: normas e procedimentos, versão 1.0. Rio de Janeiro: FGV, 1995.
- MOORE, C. W. User reactions to on-line Catalogs: an exploratory study. **College & Research Libraries**, July, p.295-302, 1981.
- PRITCHARD, S. **Library of Congress: SCORPIO user survey**. Washington, D. C., [s.n.], 1979.
- THORNE, R. & WHITLATCH, J. B. Patron on-line catalog success. **College & Research Libraries**, november, p.479-497, 1994.
- ZANAGA, M. P. Conversão retrospectiva e cooperação no processamento técnico de materiais bibliográficos: experiência no Sistema de Bibliotecas da UNICAMP. In: **SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS**, 8., 1994, Campinas. **Anais...** Campinas: UNICAMP, 1994.

### SUMMARY

The on-line catalog is the greatest impact given to users, with the possibility of a quick search of the information. Analysing the database Acervus, specifically the subject headings, its use information retrieval as in the perspective of the users as in the librarian. It has been used 97 records, and the books were presented to the judge 1 (a librarian) and to the lectures, the judges 2 for the representation of the contents. The results showed important aspects of the agreement indexes, bringing questions about the ways the subject headings in the information retrieval and in the satisfaction of the users.

**Keywords:** Subject headings - information retrieval; Subject headings - use.