



A APLICAÇÃO DA TEORIA DA CLASSIFICAÇÃO FACETADA EM BANCO DE DADOS, ATRAVÉS DA MODELAGEM CONCEITUAL

Márcio Bezerra da Silva, Dulce Amélia de Brito Neves

Resumo: Apresentamos a Teoria da Classificação Facetada (TCF) de Ranganathan como possibilidade de estruturação do conhecimento, através da organização de conceitos e da criação de relacionamentos em ambiente de Banco de Dados (BD). Aqui são elencadas as cinco categorias ou facetadas, estabelecidas por Ranganathan, chamadas de PMEST (Personalidade, Material, Energia, Espaço e Tempo). Objetiva-se, a partir de uma pesquisa aplicada e exploratória, investigar a possibilidade de aplicação da classificação facetada para organização do conhecimento, visando à recuperação da informação em BD. Também se almeja, de maneira mais específica, investigar a possibilidade da aplicabilidade da classificação facetada com a modelagem de dados em ambientes digitais. Adota-se como etapa de pesquisa o desenvolvimento do protótipo de BD, mais especificamente a fase inicial da modelagem de dados (modelagem conceitual), subsidiados por uma pesquisa bibliográfica sobre Representação Temática da Informação. O texto apresenta como resultados da pesquisa a determinação dos termos representativos aos materiais/documentos, por meio da TCF, na modelagem conceitual do protótipo de BD. Visa a organização e recuperação da informação na Instituição em pesquisa. Conclui-se que é possível adotar a TCF em ambientes digitais por meio da modelagem conceitual, além de apresentar uma estrutura de classificação facetada ilimitada, flexível e multidimensional.

Palavras-chave: Representação Temática da Informação. Teoria da Classificação Facetada de Ranganathan. Modelagem de Dados. Modelagem Conceitual.

1 INTRODUÇÃO

A humanidade, ao longo de sua história, procura organizar o cotidiano, seja para atitudes correspondentes ao dia-dia, seja para o conhecimento adquirido e desejado. Para realizar esta organização, o homem procura classificar as coisas que estão à sua volta, a exemplo das informações que circulam por todos os lados, a todo o momento.

Mesmo estando elas disponíveis em grande quantidade, o ser humano não capta todas as informações que o cercam; seleciona apenas o que considera necessário, permitindo assim a criação do que chamamos *conhecimento*.

A importância de organizar o conhecimento, através de formas de classificação, ganha cada vez mais importância. Além das classificações adotadas diariamente, há muito tempo, estudos vêm sendo desenvolvidos com o intuito de promover formas e técnicas de classificação desse conhecimento elaborado pelo homem.

Os sistemas de classificação desenvolvidos ao longo da história humana, passando pelas



filosóficas, depois bibliográficas até os esquemas de organização do conhecimento considerados mais modernos como Tesouros e Ontologias. Os sistemas de classificação dependem destes termos. Como exemplos de sistemas, citemos os mais usados no mundo, que são a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU).

Os métodos tradicionais de organização do conhecimento carecem de aperfeiçoamento. As informações contidas em ambientes digitais, principalmente na Web, encontram-se desorganizadas, dificultando a recuperação desejada; isso nos leva à ideia de que o problema não está propriamente nos dados recuperados, mas no processo anterior para isso utilizado, ou seja, na organização. Nesse sentido, em uma determinada Instituição¹, mais especificamente no Setor Pedagógico, ocorria a inexistência, ou má organização, tanto dos mínimos materiais físicos, quanto dos criados digitalmente. A partir desta situação, nos foi solicitado, pela Direção da Instituição, a criação de um banco de dados (BD), em ambiente digital, a fim de alcançar o objetivo de organizar e recuperar as informações do respectivo setor.

Partindo da premissa de que o problema se encontrava na organização, estudos foram realizados para a escolha de um sistema de classificação que fosse incorporado ao protótipo de BD, através da modelagem de dados, ou seja, que atendesse às várias necessidades da organização do conhecimento, pertinentes tanto à Ciência da Informação (CI), quanto à Ciência da Computação (CC).

Considerando esses aspectos, optamos pelo sistema de *classificação facetada* do indiano Shialy Rammarita Ranganathan (1897-1972), que foi foco na questão da organização. Também destacamos os estudos continuados, sobre a classificação facetada de Ranganathan, pelo londrino *Classification Research Group* (CRG), da escola de estudiosos como Campell, Coates, Farradane, Foskett, Milles, Vickery, Walford, entre outros. Conforme as teorias e aplicações dos antes mencionados, buscamos solucionar os problemas da *Instituição X*, respondendo a seguinte indagação, que representa a problematização do presente trabalho: De que modo a *classificação facetada* pode contribuir para organização do conhecimento registrado em BD, visando à recuperação da informação com maior valor agregado?

A busca de respostas para essa questão motivou o presente estudo a investigar a possibilidade de aplicação da *classificação facetada* para organização do conhecimento, visando à recuperação da informação em banco de dados. Especificamente: queremos investigar a possibilidade de aplicabilidade da classificação facetada com a modelagem de dados em ambientes digitais.

Calcado nas informações apresentadas, este trabalho apresenta a seguinte estrutura: no capítulo 1, contextualizamos o trabalho; no capítulo 2, abordamos a contribuição da CI, proposta sumarizada neste artigo, especialmente sobre a TCF de Ranganatham; no capítulo 3, delineamos o percurso metodológico que caracteriza a presente pesquisa; no capítulo 4, apresentamos as ações de aplicação da TCF na Modelagem Conceitual em BD, e, por fim, no capítulo 5, expomos as considerações finais e expectativas de estudos futuros.

¹ Rotularemos, ao longo do trabalho, o campo da pesquisa de Instituição X.



2 A CIÊNCIA DA INFORMAÇÃO NO PROTÓTIPO DE BANCO DE DADOS

Nos últimos tempos, como já enfatizamos, a informação ganha cada vez mais importância. Essa importância se reflete diretamente nas tomadas de decisões das pessoas, principalmente a partir da globalização, que recentemente, tem provocado a eliminação de barreiras, quando se trata de acesso ao conhecimento, fato esse nitidamente marcado pelas Tecnologias da Informação (TIs). No protótipo de BD, a informação deve passar por um processo de representação, que objetiva torná-la compreensível à linguagem dos usuários; para tal, são necessários meios, elementos, ou veículos que possibilitem o acesso a informação.

Na CI, constata-se definições que apresentam a representação como uma ação que visa “permitir a recuperação de algo, no caso, a própria informação, [...] uma vez que ela procura substituir aquilo que representa, algo que mantém informações sobre um domínio qualquer e de forma semântica” (FURGERI, 2006, p.26).

Destacamos outra definição, a de Novellino (1996, p.38), porque, para a autora, “a principal característica do processo de representação da informação é a substituição de uma entidade linguística longa e complexa – o texto do documento – por sua descrição abreviada”.

O ato de representar cria uma estrutura eficiente com fins de recuperação de informações. A partir do momento que as informações/documentos entram num sistema de informação, ou unidade de informação, estes passam por um tratamento, aliados a sentimentos: emoções, e razão, que estão relacionados ao cognitivo, visando facilitar a posterior recuperação do documentos.

Cada documento possui características próprias representadas por meio da Representação Descritiva da Informação, e pela Representação Temática da Informação, que buscam a posterior Recuperação da Informação. Conforme Feitosa (2006), ambas constituem o Tratamento Intelectual da Informação e podem ser executadas em qualquer sistema de informação.

As três ações supracitadas construíram os módulos do protótipo de BD, correspondendo à normas e características de cada elemento. No caso deste trabalho, destacaremos a Representação Temática da Informação, também chamada de *análise da informação*, ou simplesmente *indexação*; ela adota elementos simbólicos para a sua realização. Esta forma de representação se refere ao conteúdo informacional dos documentos e permite a identificação do tema, ou do assunto a que se refere, através das ações de indexação, elaboração de resumos, classificação, disseminação, busca e recuperação.

Aliada a indexação, ao consultarmos definições sobre CI, é possível verificar a presença da organização dentre suas práticas. A organização do conhecimento tem como objetivo organizar a informação, a fim de que ela seja recuperada. Entretanto, diante da problemática na quantidade informacional crescente, ocorrem transformações/adaptações nas formas de organização existentes, além de exigir a criação de novos instrumentos. É exatamente esta percepção que foi adotada para a construção do protótipo: visamos a adaptação de uma teoria de classificação, a ser apresentada na subseção 2.1, destinada a organizar o conhecimento, e, conseqüentemente, permitir a efetiva recuperação.

A compreensão do raciocínio aqui exposto fica ainda mais claro, quando retomamos o pensamento



de Berners-Lee (2001): ele defende que a organização é a ação mais importante nos processos de que envolvem a representação da informação. Sobre esta questão, Cesarino (1985 p.161-162) afirma que “grande parte das falhas [nos sistemas] se deve a erros ou omissões na interpretação do conteúdo dos documentos e na percepção da demanda das pessoas a que se destina o sistema”.

Diante da colocação do autor e com ele concordando Berners-Lee (2001), afirma que o problema em questão é ocasionado por uma falha no processo de organização, que conseqüentemente, afeta o processo de recuperação da informação. Na próxima subseção apresentaremos a TCF de Ranganathan, considerada por estudiosos, como Ingwersen e Wormell (1992), um grande contribuinte do desenvolvimento dos SRIs; antes mesmo de sua construção, os princípios da organização do conhecimento, e da metodologia sobre análise facetada podem ser os elementos iniciais do processo de modelagem de uma estrutura cognitiva; isso parece ser relevante no trabalho de análise de uma determinada área do conhecimento, assim como realizamos no protótipo.

2.1 A Teoria da Classificação Facetada de Ranganathan

Até a primeira metade do século XX, vários teóricos haviam realizado estudos sobre sistemas de classificação bibliográfica. Dentre os referidos teóricos, destacamos Cutter, Dewey, Brow e Bliss. Eles receberam influência do século passado, e repetiram algumas teorias. Entretanto, um nome se destacou durante aquele período, especialmente nos estudos voltados a classificação de assunto. Esse nome foi Shiyali Ramamrita Ranganathan, considerado por muitos, nos círculos internacionais, como o precursor do método científico da Biblioteconomia. Por isso, é comum ouvir que Ranganathan é o pai da Biblioteconomia, e, conforme Dahlberg (1979), o pai da Moderna Teoria da Classificação.

Ranganathan nasceu em 1892, numa pequena aldeia no sul da Índia. Morreu em 27 de setembro de 1972. Bacharel e Mestre em Matemática, dedicou-se a leitura das obras de W. C. Berwick Sayers, que mais tarde tornaria o seu mestre nos estudos biblioteconômicos. Após muitas leituras, e influências dos ensinamentos do seu mestre, Ranganathan partiu para uma revolução no delineamento dos sistemas de classificação da época, tema sedutor para sua mente “matemática”. A partir das leituras de Sayers, Ranganathan iniciou estudos sobre o sistema de classificação que causariam uma revolução à época. Iniciando sua enorme contribuição para os processos de organização/classificação da informação, foi influenciado pelas culturas brâmane, chinesa, além da astrologia, a quem creditava ideias abrangentes sobre o universo como um todo.

Em 1925, ao retornar a Índia, assumiu a direção da Biblioteca da Universidade de Madras e iniciou o desenvolvimento do sistema idealizado, sob o aspecto de uma classificação analítico-sintética: *Colon Classification*. Além disso, Ranganathan passou a defender a importância das bibliotecas em relação à educação do seu país. Ranganathan examinou o método da Decatomia, difundido por Kant, que representa a formação de assuntos em dez divisões, e “concluiu que, em vista do crescimento prolífico multidimensional, este método não era satisfatório” (LIMA, 2004, p.79).

Como solução, Ranganathan sugeriu a adoção da Policotomia Ilimitada, isto é, formação de assuntos com número ilimitado de divisões das áreas do conhecimento. Na necessidade de explicar como seria a sua formação de assuntos, Ranganathan buscou esclarecimentos na Árvore Baniana, que representa uma espécie de figura indiana, constituída de vários troncos secundários, formados de



tempos em tempos, adicionados ao tronco original, associando, desta forma, a uma árvore de assuntos (RANGANATHAN, 1967).

Diante de todas as contribuições apresentadas por Ranganathan, sem dúvida, a obra de maior expressão foi a *Colon Classification* que, inclusive, tem sido objeto de vários estudos nos últimos tempos. Destacamos as teses das professoras Maria Luiza de Almeida Campos (2001) e Gercina Ângela Borém de O. Lima (2004).

A *Colon Classification*, traduzida para o português como Classificação de Dois Pontos, também chamada de Classificação em Facetas, TCF, ou Classificação Analítico-Sintética, foi publicada pela primeira vez em 1933; provem da constatação de Ranganathan, quanto à rigidez dos sistemas de classificação do seu período, no caso, CDD e CDU. Compreendendo a necessidade de criação de um sistema mais flexível, Ranganathan fez uso do método hipotético-dedutivo, no plano das ideias, verbal e notacional.

Na TCF, o mapeamento do conhecimento é de grande importância, tanto para um determinado período específico, como para permitir a descoberta de conexões e analogias, entre diferentes campos do conhecimento, porque facilita a recuperação a informação (SPEZIALI, 1973). A noção de facetas tem sido frequentemente apresentada como a maior contribuição teórica da CI (MANIEZ, 1999), nos estudos sobre classificação, e tem sido inspirado vários estudos nas últimas décadas (MAS et al., 2008, p.3).

O diferencial da TCF é a utilização de uma estrutura dinâmica multidimensional, com a introdução do termo *faceta*, “que ficou sendo, nos modernos estudos sobre teoria da classificação, o substituto de característica” (BARBOSA, 1969, p.16).

Faceta é “um termo genérico usado para denotar algum componente, pode ser um assunto básico ou um isolado, de um assunto composto, tendo, ainda, a função de formar renques, termos e números” (RANGANATHAN, 1967, p.88).

Barbosa (1972, p.75) define faceta “como uma lista de termos mantendo entre si as mesmas amplas relações com a classe que lhes deu origem, ou então, ‘como um conjunto de termos produzidos pela aplicação de um amplo princípio de divisão’”.

Miranda (2005, p.132) contribui para a discussão, ao citar a definição de faceta apresentada por Ranganathan na última publicação da *Colon Classification*: “um termo genérico usado para designar qualquer componente de um assunto composto”. As facetas subordinam-se a cinco itens, definidos por Ranganathan (1967, p.398) como Categorias Fundamentais, a saber: Personalidade, Matéria, Energia, Espaço e Tempo (PMEST), Estes itens são usados “[...] estritamente no contexto da disciplina de classificação. Nada têm a ver com seu emprego na Metafísica, ou na Física”.

Personalidade [P] pode ser a essência de um determinado assunto, os objetos de estudo de uma determinada disciplina, tipos, entidades etc. Tipo de biblioteca, por exemplo, pode ser uma Personalidade da Biblioteconomia.

Consideremos como manifestação da categoria Personalidade [P], as seguintes facetas: “bibliotecas, números, equações, comprimentos de ondas de irradiação, obras de engenharia, substâncias químicas, organismos e órgãos, adubos, religiões, estilos de arte, línguas, grupos sociais,



comunidades” (VICKERY, 1980, p.212).

Matéria [M] representa manifestações, complementos, substâncias, que constituem as coisas, os objetos. Esta categoria é classificada em material e propriedade. Por exemplo, na Biblioteconomia, temos livros (material) raros (propriedade).

Energia [E], ou *ação* é a manifestação de um verbo, de uma ação, como processos, técnicas, atividades etc. Por exemplo, na Biblioteconomia, serviços como catalogação, indexação e classificação representam a categoria Energia.

Espaço [S] é uma divisão geográfica, uma manifestação de lugar, onde ocorre determinado evento, como cidades, superfícies em geral entre outros. Exemplificando, temos a cidade Rio de Janeiro, para Biblioteconomia.

Tempo [T] é uma divisão cronológica, uma manifestação de ideias em determinado tempo comum, como séculos, anos, meses, dias etc. Século XXI, para Biblioteconomia, é um exemplo da categoria em questão.

Dos cinco postulados categóricos determinados por Ranganathan (1967), *Personalidade* [P] é definida como uma categoria indefinível. Para ele, é a mais importante categoria. É a que contém os termos que atribuem à classe a sua identidade no campo do conhecimento. Esta categoria “tem causado mais dificuldades e controvérsias do que qualquer outra categoria fundamental” (ANJOS, 2008, p.197).

Tentando minimizar a problemática em discussão, Ranganathan (1967) expõe que, caso a manifestação não seja determinada como espaço, energia ou matéria, a mesma pode ser considerada como uma manifestação da categoria fundamental *Personalidade* [P]. A esta ação, Ranganathan intitula de *métodos de resíduos*.

Silva e Neves (2010), ainda no contexto da Biblioteconomia, complementam as explicações sobre as categorias fundamentais, apresentando as seguintes facetas, respectivas a cada categoria: [P] representa Personalidade (personality): *bibliotecário*; [M] representa Matéria (material): *livro*; [E] representa Energia (energy): *classificação*; [S] representa Espaço (space): *biblioteca central*; [T] representa Tempo (time): *hoje*.

Das categorias do PMEST surgem as facetas. Segundo Ranganathan (1967), categorias são as classes mais gerais, que podem ser formadas e empregadas para reunir outros conceitos. Cada faceta é formada por um número mínimo de subdivisões, denominadas *focos*; cada uma é acompanhada por algarismos arábicos decimais, constituindo seu símbolo de classificação. Após a faceta principal, surgem os *termos* que, segundo Ranganathan (1967), são as representações verbais dos conceitos em uma linguagem natural. Com base nos termos e na necessidade de ramificações de seu assunto, termos as subfacetas.

Ranganathan, além de determinar os elementos que caracterizam a TCF, em especial suas facetas, também estipulou cinco caminhos para a formação dos assuntos, e regras utilizadas para ordenar as facetas, e os elementos interiores de uma estrutura de classificação.

A Espiral do Método Científico possibilitou a Ranganathan medir o crescimento da Espiral



do desenvolvimento de novos assuntos, de maneira cíclica, identificando os estágios de Pesquisa Fundamental (*fundamental research*), Pesquisa Aplicada (*applied research*), Projeto Piloto (*pilot project*), Novo Mecanismo (*new mechanism*), Novo Material (*new material*), Novo Produto (*new product*), Utilização (*utilisation*) e Novos Problemas (*new problems*).

Diante do Método Científico apresentado, Ranganathan determinou cinco modos de representação temática, considerados preliminares, para a formação de assuntos e isolados, a saber: “Dissecação, Laminação, Desnudação, Reunião/Agregação e Superposição” (RANGANATHAN, 1967, p.351).

A *Dissecação (Dissecation)* propõe dividir o universo em partes coordenadas em um mesmo nível quantas vezes for necessário, criando para cada parte o seu próprio universo. Cada uma dessas partes é chamada de *Lâmina*, a qual pode representar um assunto básico, ou um isolado.

O modo da *Laminação (Lamination)* propõe-se a superposição de uma faceta com outra, constituindo-se camadas de assuntos básicos e ideias isoladas. Além disso, as ideias isoladas quando combinados, formam assuntos compostos.

Já a *Desnudação (Desnudation)*, também chamada de *Desfolhamento*, provoca uma diminuição progressiva da extensão, e um aumento da profundidade de um assunto básico, ou de uma ideia isolada, permitindo a formação de cadeias, através de resultados sucessivos do caminho/processo; pode representar o núcleo específico de um assunto básico, ou de uma ideia isolada. Esta é a forma de representação temática a ser adotada no protótipo, porque a Instituição tem interesse de especificar, ao máximo, a informação, já que busca a precisão na recuperação da informação.

Por sua vez, a *Reunião/Agregação (Loose Assemblage)* refere-se à combinação do assunto básico, ou composto com ideias isoladas, formando um assunto complexo, isto, uma ideia isolada complexa.

Por fim, a *Superposição (Superimposition)*, também chamada de *Sobreposição*, permite a conexão de duas ou mais ideias isoladas, pertencentes ao mesmo universo de ideias isoladas. A ideia isolada resultante desta superposição é chamada de *ideia isolada superposta*, ou *ideia isolada composta*. Este caminho é diferente do adotado pela Laminação, que permite a ligação entre isolados de universos diferentes.

Entender esses caminhos, ou seja, o da transformação do assunto em documentos do sistema de classificação, “é fundamental, pois desta compreensão resultará a sua atuação nos processos de organização e recuperação de informação” (CAMPOS, 2001, p.66). Além disso, os assuntos poderão ser relacionados de maneira ilimitada e multidimensional. A estudiosa Mas et al. (2008, p.3) complementa ao afirmar que:

[...] eficácia da classificação facetada é a sua capacidade de integrar diferentes dimensões de análise sobre os objetos de informação, para caracterizar e tornar mais fácil o acesso a informação fornecendo múltiplos caminhos de navegação para qualquer documento.



Determinados os assuntos, as facetas precisam ser organizadas no sistema de classificação. Neste sentido, Ranganathan criou uma série de Cânones e Princípios para organizar o conhecimento, e orientar na escolha da sequência de facetas, apresentados na terceira edição do *Prolegomena to Library Classification*, que tem como objetivo atingir o que ele mesmo denominou de *Cânone da Sequência Útil*.

Os cânones a serem abordados neste trabalho serão os voltados ao plano das ideias, organizados em Características, Sucessão de Características, Renque de Classes, Cadeia de Classes e Sequência de Filiação. Cada um dos cânones apresentados possui subdivisões, conforme pode ser observado no quadro 1:

CARACTERÍSTICA	SUCESSÃO DE CARACTERÍSTICAS	RENQUE DE CLASSES	CADEIA DE CLASSES	SEQÜÊNCIA DE FILIAÇÃO
Diferenciação	Concomitância	Exaustividade	Extensão decrescente	Classes subordinadas
Relevância	Sucessão relevante	Exclusividade	Modulação	Classes coordenadas
Verificabilidade	Sucessão consistente	Seqüência útil	x	x
Permanência	X	Seqüência consistente	x	x

Quadro 1: Cânones para o Trabalho no Plano das Ideias.

Fonte: Gomes, Motta e Campos (2006).

Após as informações apresentadas, podemos entender que a classificação facetada é um sistema de agrupamento por termos estruturados, por meio da análise de assunto. Diante desta análise, ocorre a identificação de facetas, ou seja, dos diferentes aspectos nele contidos, que oferecem as condições necessárias para a construção de uma estrutura organizacional, dispensando a necessidade de outros modelos. Desta forma, apresentaremos o caminho percorrido para realizar a pesquisa e quais foram as contribuições da TCF no desenvolvimento do protótipo de BD.

3 PERCURSO METODOLÓGICO

Esta pesquisa visou aplicar as teorias da classificação da CI, em ambientes digitais, da CC. Trata-se de uma Pesquisa Aplicada, definida por Oliveira (2001, p.123) como um modelo de pesquisa



que necessita de “[...] teorias ou leis mais amplas como ponto de partida, e tem como objetivo pesquisar, comprovar ou rejeitar hipóteses sugeridas pelos modelos teóricos e fazer a sua aplicação às diferentes necessidades humanas”. Implica-se também uma Pesquisa Exploratória, porque, ainda conforme o autor, a ênfase é “[...] dada à descoberta de práticas ou diretrizes que precisam modificar-se e na elaboração de alternativas que possam ser substituídas” (OLIVEIRA, 2001, p.134).

Apesquisa postula maior familiaridade com o problema, visado torná-lo explícito, ou a construir hipóteses. Envolve levantamento bibliográfico, entrevistas com pessoas que tiveram experiências práticas com o problema pesquisado e análise de exemplos que estimulem a compreensão (GIL, 1994).

Na *Instituição X*, o trabalho concentrou-se no Setor Pedagógico, responsável pela coordenação dos cursos, além da aquisição e produção dos materiais didáticos. Diante da problemática na organização desses materiais, muitos foram adquiridos mais de uma vez, como livros e programas de computador. Em vários momentos, apostilas eram elaboradas também mais de uma vez, bastando mudar o professor, de determinado curso, para que uma nova apostila, com o mesmo conteúdo, diagramação etc., fosse elaborada.

Diante da realidade do campo de pesquisa, este trabalho tratará do desenvolvimento de um protótipo de BD, intitulado *Sistema Facetado*, por meio da modelagem de dados. De acordo com Miranda (2005, p.72), “esta técnica consiste em se estabelecer um modelo de entrada e tratamento de dados a serem armazenados num sistema de informação”. Em outras palavras, “é um modo de estruturar logicamente as informações” (FURTADO; SANTOS, 1980, p.37). A modelagem de dados é constituída das etapas conceitual, lógica e física, sendo a primeira responsável pela adoção da TCF no *Sistema Facetado*.

Os estudos sobre a TCF de Ranganathan constituiu um dos objetos privilegiados desta pesquisa por acreditarmos que realizar a distribuição da informação em facetas seria uma ação positiva, já que são inerentes ao assunto ou campo principal. De acordo com as palavras de Silva e Neves (2010, p.14-15), “no interior de cada faceta, os termos que as constituem estão aptos a novos agrupamentos pela aplicação de outras características ou classes divisionais, dando origem às facetas menores, que chamaremos de subfacetas”.

A maior contribuição da CI para este estudo se deve à TCF, apresentando o nome sugestivo ao protótipo de BD. De fato, o que propiciou a escolha da TCF para este trabalho foi exatamente a flexibilidade, ou seja, a possibilidade de criar e recriar novos agrupamentos entre as facetas, sem alterar significativamente a estrutura de programação e relacionamentos das tabelas do BD.

Contextualizado os elementos que constituem o *Sistema Facetado*, apresentaremos o resultado de sua construção, na próxima seção, conforme o nível inicial da modelagem de dados: modelagem conceitual.

4 APLICANDO A TCF NA MODELAGEM CONCEITUAL EM BANCO DE DADOS



O protótipo de BD promulgou a representação da realidade do Setor Pedagógico da *Instituição X*, que necessitava recuperar informação. Para realizá-la, foi usada a modelagem conceitual, definida como “a representação abstrata e simplificada de um sistema real, com a qual se pode explicar, ou testar o seu comportamento, no todo ou em partes” (COUGO, 1997, p.7).

Na modelagem conceitual não são levadas em consideração atribuições computacionais relacionadas à escolha, neste caso, do Sistema Gerenciador de Banco de Dados (SGBD), do Sistema de Banco de Dados (SBD) e do formato de estruturação do sistema. Neste momento, na tentativa de representar a realidade da *Instituição X*, a técnica de indexação e os estudos da TCF foram de fundamental importância, já que, a partir dos subsídios da CI supracitados, os termos representativos da Instituição foram adquiridos conforme a indexação por extração e, em seguida, organizada no *Sistema Facetado*.

Segundo Machado e Abreu (1996, p.21), na modelagem conceitual, o analista “[...] deve se concentrar na observação dos fatos relevantes que ocorrem na realidade, com a finalidade de construir um sistema que possa automatizar as necessidades de informação da mesma”.

Tomando partido das compreensões dos autores, no protótipo deste trabalho, adaptamos a estrutura conceitual do Mapa Hipertextual (MHTX) criado por Lima (2004), em sua tese, acrescido de algumas ações, a saber: conversar com os possíveis usuários do sistema; compreender as atividades realizadas pelos usuários na Instituição; identificação dos materiais pertencentes ao Setor Pedagógico; leitura técnica dos materiais para a extração dos termos representativos; análise facetada; organização de todos os termos selecionados e suas respectivas relações.

A análise facetada permitiu incorporar propriedades adicionais às informações, sejam elas textuais ou multimídia, além de fornecer dados para a análise das categorias de dados presentes em ambientes digitais, como é o caso do protótipo de BD, afetando, assim, a sintaxe potencial do referido sistema. Porém, para adicionar novas propriedades a tais elementos e alterar a leitura da sintaxe do ambiente digital em discussão, foi necessário sua análise, a fim de levantar e definir a terminologia do assunto para que posteriormente os termos presentes no sistema fossem analisados e distribuídos em facetas.

Na formação dos assuntos, dentre os métodos “Dissecação, Laminação, Desnudação, Reunião/Agregação e Superposição” (RANGANATHAN, 1967, p.351), apresentados na subseção 2.1, adotamos a *Dissecação* e a *Desnudação*, já que os termos extraídos da indexação foram mantidos na linguagem natural, ou seja, conforme são apresentados nos materiais, não passando pela tradução de qualquer vocabulário controlado. Além disso, também nos baseamos na proposta do MHTX criado por Lima (2004).

A *Dissecação* (*Dissection*) permitiu dividir o universo dos materiais, e identificar que “Educação Profissionalizante em Informática” era o assunto básico, também rotulado pela Instituição de “Capacitação Profissional”, ou “Capacitação Profissionalizante”.

O outro método utilizado, a *Desnudação* (*Desnudation*), também permitiu representar o núcleo



específico de um assunto básico, ou de uma ideia isolada dos materiais consultados. A *Desnudação* foi considerada o método mais importante para o sistema, pois a Instituição necessitava que a recuperação de informações fosse precisa, exigindo maior especificidade dos assuntos, de acordo com a exemplificação do quadro 2:

ASSUNTO BÁSICO
■ Cursos
■ ■ Curso Básico
■ ■ ■ Curso de Introdução a Microinformática

Quadro 2: Exemplo de Desnudação no Sistema Facetado.
Fonte: Dados da pesquisa (2011).

Para realizar os métodos apresentados, foi necessário executar algumas ações e, baseado nas orientações apresentadas por Lima (2004) em sua tese, adotamos a leitura minuciosa do documento; identificação dos conceitos representativos ao conteúdo semântico do documento; seleção dos termos mais relevantes ao assunto básico; seleção dos termos com um conteúdo semântico significativo; e representação natural dos conceitos selecionados, ou seja, não utilizando um vocabulário controlado.

Apesar de as orientações serem construídas por Lima, com olhar nos *hiperlinks*, verificamos aplicação de ambas na modelagem conceitual no protótipo de BD, constituindo a primeira ação da CI utilizada no sistema: a indexação/análise facetada. Segundo esta ação de representação da informação, apresentamos uma parcela dos termos (quadro 3) necessários ao desenvolvimento de uma estrutura, chamada neste trabalho de *Mapa Categorical do Sistema Facetado*, contendo como classe geral *Educação Profissionalizante em Informática*.

Administrativo	Manhã
Aluno	Modular
Apostila	Nivelamento
Apresentação	Palestras
Artes	Primeiro
Cadastro	Professor
Capacitação	Profissionais
Cartões	Propriedade
CD-ROM	Relatórios
Cursos	Segundo
Documentos	Semestre



Internet	Pedagógico
Interdisciplinar	<i>Software</i>
Microinformática	Tarde
Laboratório	Técnico
Livro	Turno

Quadro 3: Lista de termos para o Sistema Facetado.

Fonte: Dados da pesquisa (2011).

Determinados os assuntos, estes precisavam ser organizados conforme suas características. A organização ocorreu segundo as cinco categorias fundamentais de Ranganathan (1967): Personalidade, Matéria, Energia, Espaço e Tempo (PMEST). Na figura 1, temos um trecho da organização estruturada pela categoria fundamental Tempo [T], segundo o plano notacional:

5. Tempo
a. Turno
i. Manhã
ii. Tarde
iii. Noite
b. Semestre
i. Primeiro
ii. Segundo

Figura 1: Organização notacional estruturada pela categoria Tempo [T].

Fonte: Dados da pesquisa (2011).

O plano notacional fez parte da construção do sistema, contribuiu para a organização dos assuntos, entretanto, diante do seu acervo ser constituído em sua maioria por documentos digitais, sendo os impressos digitalizados futuramente, o Setor Pedagógico descartou a utilização da notação da *Colon Classification*, CDD ou CDU. A notação ocorrida foi adaptada conforme o citado setor, ou seja, a cada entrada, os documentos recebem um código chamado de ID². Sendo assim, a notação, criada no módulo de classificação do sistema, é formada pelo símbolo do assunto, neste caso “INF”, por tratarmos exclusivamente de documentos dos cursos de informática, adicionado ao número de ID do documento cadastrado. Por exemplo, o livro “Introdução a Microinformática” foi o primeiro livro cadastrado, recebendo a “ID 1”. Ao associarmos a ID ao seu respectivo assunto, temos a notação resultante “INF1” e assim sucessivamente.

Baseado no plano das ideias, complementamos a estrutura de organização do sistema adotando a proposta do CRG que usa no princípio das categorias fundamentais, categorias baseadas no contexto do próprio assunto. Segundo Lima (2002, p.116), o CRG sugere “[...] que nenhuma lista pode ser imposta mecanicamente ao assunto. Além disso, o CRG acredita que nenhuma lista deva

² ID: Identidade gerada pela numeração automática do sistema, conforme a entrada do material no protótipo.



ser necessariamente exaustiva ou aplicada a todos os assuntos”. Neste quesito, tal adoção foi muito importante para este trabalho, diante da contribuição dos aspectos que envolvem a passagem de uma organização estruturada notacional para uma estrutura sistemática de termos (figura 2).

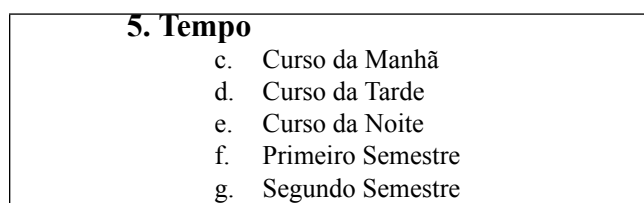


Figura 2: Organização sistemática estruturada pela categoria Tempo [T].

Fonte: Dados da pesquisa (2011).

Entretanto, alguns conceitos receberam outras denominações para representar, de maneira mais fidedigna, a realidade do Setor Pedagógico. Assim, Personalidade [P] assumiu Tipologia; Matéria [M] assumiu Propriedade (qualidade) e Matéria (substância); Energia [E] assumiu Ação; Espaço [S] assumiu Lugar; e Tempo [T] manteve sua rotulação.

Para a estrutura do sistema, as categorias e suas respectivas facetas precisavam de uma ordem de apresentação, ou seja, de uma *Ordem de Citação*, que é a ordem de aplicação dos princípios de divisão. O exemplo da figura 1, para ser apresentado desta maneira, foi estruturado segundo regras que objetivavam a *Ordem de Citação* deste trabalho.

Ranganathan, para organizar o conhecimento e orientar na escolha da sequência de facetas, adotou como referência a *Policotomia Ilimitada* e a *Árvore Baniana*. Sua ordenação se deu pela criação de uma série de *Canônes* e *Princípios*, cujo objetivo é atingir o que ele mesmo denominou de *Canône da Sequência Útil*.

Buscando adotar uma Sequência Útil para este trabalho, cinco regras, entre cânones e princípios foram seguidas: *Cânone Cadeia de Classes: Extensão Decrescente*; *Princípio da Contiguidade Espacial*; *Cânone Renque de Classes: Seqüência Útil* e *Sequência Consistente*; e *Princípio da Seqüência Canônica*.

No *Cânone Cadeia de Classes*, aplicamos a regra *Extensão Decrescente*, que apresenta a classe geral antes das específicas. No sistema, observamos o aprofundamento em uma das sequências da categoria Personalidade [P], iniciando no mais geral “Profissionais”, passando por “Professor” e decrescendo até o mais específico “Professor de Informática”.

A segunda regra adotada para organizar a sequência de facetas foi *Princípio do Posterior-na-Evolução*, que segue a organização de conceitos ligados a processos evolutivos. Esta regra pode ser observada em uma sequência da categoria Matéria [M], no foco “Nivelamento”, ao exigir que o aluno seja aprovado em diferentes módulos para adquirir conhecimentos aprofundados de determinado curso: Básico, Intermediário e Avançado.

No *Cânone Renque de Classes*, adotamos a regra *Seqüência Útil*, já que seguimos sequências



de classes consideradas úteis aos propósitos da classificação das informações no sistema. Como exemplo, temos, na categoria Personalidade [P], no foco “Documentos”, uma sequência de renques iniciada por “Administrativo” e na especificação “Cadastro de Professor” e “Cadastro de Aluno”. Esta é uma sequência útil para o sistema, baseada nas atividades realizadas na Instituição em pesquisa.

Ainda no *Cânone Renque de Classes*, adotamos também a regra *Sequência Consistente*, porque pode ocorrer de conceitos iguais surgirem em mais de um ponto da estrutura de classificação, exigindo que todos tenham a mesma ordenação. Na categoria Personalidade [P], foco “Documentos”, temos em “Ensino”, uma sequência de facetas formada por “Apostila”, “Livro” e “CD-ROM”. Já na categoria Matéria [M], foco “Por Matéria (substância)”, temos a mesma sequência formada por “Apostila”, “Livro” e “CD-ROM”. Na primeira categoria, a sequência refere-se à tipologia de documentos utilizados no “Ensino”; na segunda categoria, a sequência está associada à substância que formam os materiais: como papel e plástico.

Outra regra seguida foi o *Princípio da Sequência Canônica*, que consagra a tradição, ou seja, a ordem mais conhecida comumente por todos os que usam o sistema. Neste caso, na categoria Tempo [T]; para o foco “Turno”, compreendemos que a sequência conhecida por todos é “Manhã”, “Tarde” e “Noite”.

Definida a modelagem conceitual, em outro trabalho será apresentada a representação da modelagem lógica e física, em cujas etapas da modelagem de dados, as informações deverão ser organizadas e relacionadas como subsídios para a implementação do *Sistema Facetado* no Setor Pedagógico.

5 CONSIDERAÇÕES FINAIS

Ao longo do presente trabalho, foi possível perceber a importância da organização para a efetiva recuperação da informação. Para muitos, o problema dos resultados de busca se refere às ferramentas de busca, ou ao excesso de informações. Entretanto, buscando autores que defendem outra perspectiva, nos baseamos naqueles que asseveram que o problema se encontra na organização, ou seja, como as coisas, e os objetos são classificados.

Nas aplicações desenvolvidas pelos profissionais de Ciência da Computação (CC) e áreas afins, é nítida a não preocupação com os elementos que estão por trás da interface, por trás da máscara do sistema visualizada pelos usuários. Em várias discussões e fóruns, verificamos a preocupação pelo desenvolvimento de ferramentas de busca cada vez melhores, em plataformas que ofereçam um visual deslumbrante e amigável.

Baseado nos estudos teóricos buscou-se discutir o que se encontra por trás da interface do sistema, ou seja, a organização/classificação, preocupação esta respalda na Ciência da Informação (CI). O que propormos a fazer é o chamado “*trabalho sujo*”, isto é, um trabalho que exige ter conhecimento sobre representações e sistemas de classificação.



A CI nos ofereceu os elementos necessários para a construção de um sistema para a catalogação, classificação e, por fim, recuperação dos materiais do Setor Pedagógico da *Instituição X*. Neste certame, ao longo dos estudos para a definição dos métodos de organização do conhecimento a serem usados, identificamos a necessidade de aperfeiçoamento dos métodos tradicionais, especialmente quando falamos da Web, onde as informações encontram-se desorganizadas. Foi exatamente esta constatação que nos levou a crer que o problema não está na recuperação, mas no processo de organização.

Nas práticas da CI, enraizadas na Biblioteconomia, mais especificamente nas bibliotecas tradicionais, encontramos sistemas de classificação chamados bibliográficos, originalmente criados e usados em coleções de livros. Entretanto, características desses sistemas, como a organicidade de assuntos e as relações entre eles são plausíveis de aplicação em ambientes digitais. Adotamos a Teoria da Classificação Facetada (TCF), do matemático e indiano Shiyali Ramamrita Ranganathan (1892-1972), complementada pelo princípio das categorias fundamentais, proposta pelo *Classification Research Group* (CRG). Após os devidos estudos, ambos foram inseridos em um ambiente digital, neste caso, um protótipo de Banco de Dados (BD).

Ao longo da idealização do sistema, muitas foram às semelhanças encontradas. Dentre elas, percebemos que, assim como na classificação facetada, os sistemas computacionais por modelagem de dados, em BD, também objetivam a estruturação do conhecimento, através da organização de seus conceitos e da criação de relacionamentos entre eles, permitindo o mapeamento de uma área de assunto e a inclusão de novos conceitos, sem alterar a estrutura do sistema.

Segundo as semelhanças entre a TCF de Ranganathan e os BDs, foi iniciado o desenvolvimento de um Sistema de Banco de Dados (SBD). Este protótipo de BD, que intitulamos *Sistema Facetado*, tem como objetivo solucionar problemas, tipo: produção/confecção da mesma apostila várias vezes, e a não localização de suportes como fitas VHS, CDs etc., devido à inexistência, ou má organização ali presente, tanto nos mínimos materiais físicos, como nos criados digitalmente.

Concluída a aplicação da TCF no *Sistema Facetado*, por meio da modelagem conceitual, constatamos que a análise de determinado assunto por facetadas (PMEST), significou que o mesmo era visualizado pelas manifestações das características, ou facetadas a ele atribuídas, assunto que, no caso deste trabalho, chamamos de conceitos. Estas manifestações foram muito importantes, pois o *Sistema Facetado* tornou-se multidimensional e ilimitado, representados pelas facetadas e organizadas por categorias, ou seja, uma lista de conceitos, hierarquizados igualmente aos seus respectivos assuntos gerais, refletindo a aplicação de um princípio básico de divisão.

Muitas foram às contribuições da TCF para o *Sistema Facetado*, o que comprova sua aplicação em ambientes digitais. Com isso, desejamos o surgimento de mais estudos sobre esta temática, além de acreditar que o presente trabalho inspire aqueles que decidam enveredar por esse caminho, assim como decidi, graças às leituras encorajadoras das Teses das professoras Maria Luíza de Almeida



Campos e Gercina Ângela Borém de O. Lima.

O *Sistema Facetado* é um protótipo. Muito resta ainda a fazer. Este trabalho apresentou parte da modelagem de dados: a modelagem conceitual. Em outro momento apresentaremos a modelagem lógica, e física, além do *Sistema Facetado* em funcionamento. Ainda faz-se necessário salientar: expandiremos o uso do sistema para os materiais dos demais cursos oferecidos pela Instituição.

Abstract: The present study focalizes Ranganathan's Theory of Faceted Classification (TFC) as possibility of knowledge structuring through the organization of concepts and relationships creation in a environment of Database (DB). Contributions of Information Science (IS) for the research by defend the importance of organization for effective information retrieval is discussed: descriptive and thematic representation of information. The five categories established by Ranganathan and facets in this system called PMEST are cited: Person, Material, Energy, Space and Time. From an applied and exploratory research, the object is to investigate the possibility of applying the faceted classification to knowledge organization, looking for information retrieval in the database. More specifically, we also want to check the applicability of faceted classification with data modeling in digital environments. Use as the development stage of the research prototype database, specifically the initial phase of data modeling (conceptual modeling), subsidized by a literature search on Thematic Representation of Information. Presents as search results the determination of the terms representing the materials/documents, through the TFC in conceptual modeling of the prototype database, searching the organizing and retrieving information in the institution researched. We conclude that it is possible to adopt the TFC in digital environments and that this action may be taken by the conceptual model, featuring a faceted classification structure unlimited, flexible, and multidimensional.

Keywords: Thematic Representation. Ranganathan's Theory of Faceted Classification. Data Modeling. Conceptual Modeling.

REFERÊNCIAS

ANJOS, L. dos. **Sistemas de classificação do conhecimento na Filosofia e na Biblioteconomia:** uma visão histórico-conceitual crítica com enfoque nos conceitos de *classe*, de *categoria* e de *faceta*. São Paulo, 2008. 290f. Tese (Doutorado em Ciência da Informação) – Curso de Pós Graduação em Ciência da Informação – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2008.

BARBOSA, A. P. Classificações facetadas. **Ci. Inf.** Rio de Janeiro, v.1, n. 2, p. 73-81, 1972. Disponível em: <revista.ibict.br/index.php/ciinf/article/download/1665/1271>. Acesso em 11 jan. 2010.

_____. **Teoria e prática dos sistemas de classificação bibliográfica.** Rio de Janeiro: Instituto Brasileiro de Bibliografia e Documentação, 1969.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic WEB. **Scientific American**, New York, v. 284, n. 5, p.35-43, may 2001. Disponível em: <<http://www.scientificamerican.com/2001/0501issue/0501bernerslee.html>>. Acesso em: 12 jan. 2010.

CAMPOS, M. L. A. **A organização de unidades do conhecimento em hiperdocumentos:** o modelo conceitual como um espaço comunicacional para realização da autoria. 2001. 190f. Tese (Doutorado em Ciência da Informação) - CNPq/IBICT-URFJ/ECO, Rio de Janeiro, 2001.



CESARINO, M. A. da N. Sistemas de recuperação da informação. **Revista da Escola de Biblioteconomia** da UFMG. v. 14, n. 2, p.157-68, set. 1985. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/reb/>>. Acesso em: 29 jan. 2010.

COUGO, P. **Modelagem conceitual e projeto de banco de dados**. Rio de Janeiro: Campus, 1997.

DUARTE, E. A.; CERQUEIRA, R. F. P. de. Análise Facetada: um olhar face a modelagem conceitual. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 4, n. 2, p. 39-52, jan./jun. 2007. Disponível em: <http://www.sbu.unicamp.br/seer/ojs/index.php/sbu_rci/article/viewFile/339/221>. Acesso em: 02 jan. 2010.

FURGERI, S. **Representação de informação e conhecimento**: estudo das diferentes abordagens entre a ciência da informação e a ciência da computação. Campinas: PUC-Campinas, 2006.

FURTADO, A. L.; SANTOS, C. S. dos. **Organização de Banco de Dados**. 2. ed. Rio de Janeiro: Campus, 1980.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 4. ed. São Paulo: Atlas, 1994.

GOMES, H. E.; MOTTA, D. F. da; CAMPOS, M. L. de A. **Revisitando Ranganathan**: a classificação na rede. 2006. Disponível em: <<http://www.conexao rio.com/bit i/revisitando/revi sitando.htm#canones>>. Acesso em: 10 jan. 2011.

INGWERSEN, P.; WORMELL, I. Ranganathan in the perspective of advance information retrieval. **Libri**, v. 42, n. 3, p.184-201, jul./set. 1992.

LIMA, G. A. B. A análise facetada na modelagem conceitual de sistemas de hipertexto: uma revisão de literatura. **Perspectiva em Ciência da Informação**. Belo Horizonte, v. 7, n. 2, p. 189-196, jul./dez. 2002. Disponível em: <portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/407/219>. Acesso em 20 jan. 2010.

_____. **Mapa Hipertextual (MHTX)**: um modelo para organização hipertextual de documentos. 2004, 199f. Tese (doutorado em Ciência da Informação). Curso de Pós Graduação em Ciência da Informação – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Minas Gerais, 2004.

MACHADO, F.; ABREU, M. **Projeto de Banco de Dados**: uma visão prática. 11. ed. São Paulo: Editora Érica, 1996.

MANIEZ, J. Des classifications aux thésaurus: du bon usage des facettes. **Documentaliste - Sciences de l'information**, v. 36, n. 4-5, p.249-262, 1999. Disponível em: <<http://www.adbs.fr/des-classifications-aux-thesaurus-du-bon-usage-des-facettes-13338.htm?RH=REVUE>>. Acesso em: 12 jan. 2011.

MAS, S. [et al.]. Classification à facettes et modèles à base de points de vue: Différences et complémentarité. **36e congrès annuel de ACSI**: University of British Columbia, Vancouver, 5-7 jun. 2008. Disponível em: <http://www.caisacsi.ca/proceedings/2008/mas_2008.pdf>. Acesso em: 24 out. 2010.

MIRANDA, M. L. C. de. **Organização e Representação do Conhecimento: fundamentos teórico-metodológico na busca e recuperação da informação em ambientes virtuais**. 2005, 353f. Tese (Doutorado em Ciência da Informação). Curso de Pós Graduação em Ciência da Informação –



ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO Brasília, Distrito Federal 23 a 26 de outubro de 2011

Convênio CNPQ/IBICT – UFRJ/ECO, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

OLIVEIRA, S. L. de. **Tratado de Metodologia Científica**: Projetos de Pesquisas, TGI, TCC, Monografias, Dissertações e Teses. São Paulo: Pioneira, 2001.

RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publ. House, 1967.

SILVA, M. B. da. NEVES, D. A. de B. Estudo sobre o uso da teoria da classificação facetada em banco de dados. XI Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), Rio de Janeiro, 25-28 de outubro de 2010. **Anais**. Disponível em: <<http://congresso.ibict.br/index.php/enancib/xienancib>>. Acesso em: 05 jan. 2011.

SPEZIALI, P. Classifications of the sciences. **DICTIONARY of the history of ideas**. New York: Scribners, 1973. p. 462-467.

VICKERY, B. C. **Classificação e indexação nas ciências**. Rio de Janeiro: BNG/Brasilart, 1980.