

GT 2: Organização e Representação do Conhecimento

**PROTOTIPAGEM DE BANCO DE DADOS: o uso da teoria da classificação
facetada na modelagem de dados**

Comunicação Oral

Márcio Bezerra da Silva – UFPB

Dulce Amélia de Brito Neves – UFPB

m_informatica@hotmail.com

Resumo: Apresenta a inserção da Teoria da Classificação Facetada (TCF) de Ranganathan na prototipagem de Banco de Dados (BD) a partir da modelagem de dados, mais especificamente nas etapas de programação: lógica e física. Discute modelagem de dados em BD como técnica que deverá representar a realidade do ambiente que o sistema será inserido. Objetivou-se um percurso metodológico formado pela Pesquisa Aplicada, Exploratória e Bibliográfica, de abordagem qualitativa, apresentar a aplicação da *classificação facetada* para organização do conhecimento, visando à recuperação da informação em banco de dados, como também, investigar a possibilidade de aplicabilidade da *classificação facetada* com a modelagem de dados em ambientes digitais e expor um panorama de uso do protótipo de BD. Coleta, como resultados da pesquisa, as contribuições da TCF na modelagem lógica e física, respeitando os métodos de formação de assunto de Ranganathan, Dissecação e Desnudação, feito na modelagem conceitual (outro trabalho). Realiza, na modelagem física, a escolha de recursos de *software*; a confecção dos módulos do sistema em Cadastro (catalogação), Classificação e Pesquisa (busca); e contribuições da TCF na facetagem ilimitada e filtragem nas ocorrências de buscas, tornando as pesquisas mais precisas. Conclui-se que os subsídios para o desenvolvimento do *Sistema Facetado* foi a TCF e a modelagem de dados em BD. Finaliza almejando o desejo de outros estudos, outros modelos, os quais poderão nos ajudar a ampliar esta discussão para ambientes mais gerais, abrangentes, como é o caso da Web.

Palavras-chave: Teoria da Classificação Facetada de Ranganathan. Modelagem de Dados. Modelagem Lógica. Modelagem Física.

1 INTRODUÇÃO

A cada instante, o homem procura organizar o que está a sua volta, segundo alguma forma de classificação. Essa forma de classificação segue critérios, características, que o permitirão enxergar o mundo e se inserir em grupos. Para tanto, a informação será o elemento condutor da realização de tais ações, formada por dados que necessitam de organização, estruturação e representação, conforme signos e símbolos.

Segundo o conjunto dessas informações, o homem monta, elabora o seu conhecimento, necessário para a sua sobrevivência, num processo de constante mudança. Para Da Silva e Neves (2011, p. 540, *grifo dos autores*), “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*”. Robredo (2003) complementa a discussão ao afirmar que o conhecimento está relacionado à maneira em que a informação é usada na formação ou transformação de algo.

Na prerrogativa apresentada, estudos são realizados para aprimorar e/ou criar instrumentos de organização do conhecimento, e que estes devem acompanhar as mudanças impostas pelo atual fluxo de informação, envolto de adventos tecnológicos. Com isso, destacamos a Ciência da Informação (CI), considerada uma área de estudo interdisciplinar e que surgiu por volta da Segunda Guerra Mundial, devido ao aumento na produção informacional provocado pelo avanço técnico-científico. É uma ciência formada por áreas, conforme Saracevic (1996), como Biblioteconomia, Matemática, Lógica, Lingüística, Psicologia, Ciências Cognitivas, Ciência da Computação (CC), Comunicação entre outras.

Diante da presença de muitas áreas do conhecimento na constituição da CI, sobressaímos a Biblioteconomia na perspectiva da classificação bibliográfica; e a CC, no desenvolvimento e implementação de Banco de Dados (BD). Ambas os campos de estudos nos ofereceram subsídios na busca por soluções aos problemas de uma determinada Instituição¹, mais especificamente em seu Setor Pedagógico, quanto à inexistência ou má organização, tanto dos mínimos materiais físicos, quanto dos criados digitalmente.

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

Compreendido a importância da organização nesse processo, criamos um BD em solicitação a Direção da Instituição, com o objetivo de organizar e recuperar as informações do respectivo setor. Neste certame, no âmbito dos estudos sobre classificação bibliográfica, escolhemos a *Teoria da Classificação Facetada* (TCF) do indiano Shialy Rammarita Ranganathan (1897-1972), onde o sistema adotado foi incorporado ao protótipo de BD, através da modelagem de dados da CC. Inicialmente realizamos um estudo teórico sobre o uso da TCF em BD para determinar os elementos básicos e etapas para a construção do sistema, sendo tema do XI Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB); em seguida apresentamos a fase inicial da modelagem de dados, a etapa conceitual, discutido no XII ENANCIB; e na atuação edição temos à pretensão de apresentar as outras etapas da modelagem de dados (modelagens lógica e física), nos permitindo observar o protótipo de BD em funcionamento.

A referida observação calca-se na contínua busca em responder a seguinte indagação: a *classificação facetada* pode contribuir para organização do conhecimento registrado em BD, visando à recuperação da informação com maior valor agregado?

A partir do contexto apresentado, baseado em um percurso metodológico formado pela Pesquisa Aplicada, Exploratória e Bibliográfica, de abordagem de coleta de dados qualitativa, almejamos de maneira geral apresentar a aplicação da *classificação facetada* para organização do conhecimento, visando à recuperação da informação em banco de dados. De forma específica, buscamos investigar a possibilidade de aplicabilidade da *classificação facetada* com a modelagem de dados em ambientes digitais e expor um panorama de uso do protótipo de BD.

2 MODELAGEM DE DADOS EM BD

O BD deve ser criado respeitando a realidade do ambiente que será inserido. Para tal, fatores como os usuários, tipos de informação, recursos de *hardware* e *software* adotados entre outros, devem ser levados em consideração. Em outras palavras, será “[...] um modo de estruturar logicamente as informações” (FURTADO; SANTOS, 1980, p. 37).

Para Elmasri e Navathe (1994), a modelagem ou modelo de dados é um conjunto de conceitos que se usa para descrever a estrutura do BD e certas restrições que o banco deve garantir. Conforme Miranda (2005), a estruturação por modelo de dados permite

ao usuário recuperar dados de forma rápida e eficiente, cujo objetivo é incluir dados em uma estrutura que possibilite transformar os dados originais em vários tipos de saídas como formulários, relatórios, etiquetas ou gráficos.

Um dos pontos críticos da modelagem de dados está além do conhecimento técnico, como enfatizamos anteriormente. Estamos nos referindo ao conhecimento sobre a realidade em que o BD será construído, buscando solucionar algum tipo de problema. Neste sentido, Machado e Abreu (1996, p. 22) defendem a preocupação com a realidade ao expor que:

Para registramos as necessidades de informação de uma realidade, precisamos fazer uso de um modelo, ou seja, algo que nos mostre como as informações estão relacionadas (fatos). E com base no modelo criado, os analistas podem interagir com os usuários validando seus objetivos e metas, permitindo a construção de um sistema de informações cada vez mais próximo de realidade.

Aliados a preocupação com a realidade, na modelagem de dados existe uma quantidade de abstração de informações que varia conforme o tipo de usuário e modelo de dados escolhido. Chamado de *Nível de Abstração dos Dados*, esta característica representa a omissão de detalhes sobre o armazenamento das informações para alguns usuários (KORTH; SILBERSCHATZ, 1995) e classifica-se em três modelos.

O *Nível Conceitual* é utilizado pelos administradores do BD e define, de forma independente, a implementação do BD, quais os dados que serão armazenados e como será o relacionamento entre eles, bem como os níveis de restrição. Neste nível são adquiridos, do mundo real, os conceitos/dados que farão parte do BD, objetivando solucionar algum problema, preocupação essa defendida por Machado.

O *Nível Lógico* define como o BD será implementado. Neste momento, tem-se uma visão mais detalhada sobre o armazenamento dos dados. Além disso, é imprescindível conhecer/decidir o tipo de BD que será implementado (rede, hierárquico, relacional e orientado a objeto).

Terceira e última característica, o *Nível Físico*, descreve a forma como os dados são armazenados e os aspectos físicos de implementação, ou seja, preparar o BD para a prática. Neste nível ocorre à escolha dos recursos de *software* envolvidos na sua elaboração, variando a metodologia de aplicação lógica conforme os recursos oferecidos pelo SGBD.

A modelagem de dados é a principal característica utilizada para se classificar os SGBDs, sendo os mais recentes baseados em modelos de Programação Orientada a

Objetos (POO). Segundo Ricarte (2001), POO “pressupõe uma organização de software em termos de coleção de objetos discretos incorporando estrutura e comportamento próprios”.

A organização/classificação está além de uma técnica limitada para arranjar livros em estantes de bibliotecas, especialmente pelo nosso desejo de adotar a TCF em BD. A referida técnica possui raízes na nossa memória, no social, e que pode ser aplicada também em ambientes digitais, assim como defende Langridge (1977, p. 105):

[...] devemos portanto ser cuidadosos para não confundir o uso de ‘classificação’ como um modo fundamental de pensamento e operação com o uso que implica um sistema totalmente fixo incluindo arranjo final em ordem classificada.

A afirmação do autor está condizente aos dias de hoje, principalmente quando tratamos da Web, um ambiente totalmente dinâmico e sem fronteiras, que se tornou um dos nossos ambientes de estudo e pesquisa, enraizado na globalização e que provoca a eliminação das barreiras para o acesso ao conhecimento produzido pela humanidade.

Com base na fundamentação apresentada até este ponto, apresentaremos a realização da modelagem de dados para a construção do protótipo de BD, respeitando as três etapas aqui mencionadas.

3 A MODELAGEM LÓGICA E FÍSICA NA PROTOTIPAGEM DO BD

Como já foi apresentada, a modelagem de dados é composta por três etapas. Na primeira, a conceitual, a maior contribuição da CI foi a TCF, apresentando o nome sugestivo de *Sistema Facetado* ao SBD. 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, que propiciou a sua escolha para este trabalho. Tudo se baseou em nossa preocupação de representar a realidade do Setor Pedagógico da *Instituição X*, que necessitava recuperar informação, e para realizá-la, Lima (2004, p.107) “nos remete [...] aos procedimentos da análise facetada, com sua característica de abrigar qualquer assunto a partir de diversos pontos de vista”. Esta interpretação se deve a sua característica multidimensional e não estruturada em tabelas rígidas, como ocorrem nos sistemas chamados *top-down*, ou seja,

conforme Da Silva (2010), a ordenação parte de cima para baixo, numa sequência imutável e inflexível.

Após a modelagem conceitual, chegamos às instruções de implementação do sistema, um dos focos deste trabalho. Na modelagem lógica, são apresentados maiores detalhes sobre o armazenamento dos dados e como serão relacionados. Conforme Heuser (2004), a modelagem lógica é o resultado ou produto da conversão de uma modelagem conceitual para um determinado tipo de banco de dados. Para Machado e Abreu (1996, p.24), “a modelagem lógica objetiva descrever “[...] as estruturas que estarão contidas no banco de dados, de acordo com as possibilidades permitidas pela abordagem, mas sem considerar, ainda, nenhuma característica específica de um [...] SGBD [...]”. O resultado é um esquema lógico de dados, construído a partir de um tipo de BD.

Dentre eles, adotamos o modelo relacional, que consiste em um conjunto de tabelas, onde suas linhas representam um relacionamento entre um grupo de valores. Atualmente, muitos dos SGBDs fazem uso do modelo relacional no desenvolvimento de suas aplicações. O autor complementa ao afirmar que “a base da moderna tecnologia de banco de dados é, sem dúvida, o modelo relacional” (DATE, 2003, p.93).

O modelo relacional, pela sua multidimensionalidade, foi o escolhido por apresentar características semelhantes à TCF, especialmente pela inserção de novos dados sem alterações significativas ao modelo estruturado. Neste sentido, foi possível inserir facetas e relaciona-las entre as tabelas do *Sistema Facetado*.

A representação da modelagem conceitual em lógica foi criada no *software Microsoft (MS) Office Visio*, versão 2007. Apresentando formas predefinidas, modelos inteligentes e exemplos, o *MS Office Visio* oferece uma grande variedade de opções para atender às suas necessidades de diagramação em TI, negócios, gerenciamento de processos e muito mais.

O *MS Visio*, versão 2007, é uma excelente ferramenta para representações em diagramas de determinada realidade. Possui vários esquemas de desenho, além de permitir que outros sejam baixados da Internet, como é o caso dos diagramas com a notação original de Peter P. Chen, um dos grandes estudiosos sobre BD relacional.

Sua escolha se deu pela fácil criação de diagramas, além da possibilidade de modificá-los sempre que houvesse necessidade, atendendo, assim, as exigências multidimensionais da TCF e dos BDs, mencionados anteriormente. O *Visio* possui uma categoria específica para BD e, dentre as suas opções, usamos o Diagrama de Modelo

de Banco de Dados (figura 1). Quanto aos benefícios oferecidos na construção de diagramas, citemos o adimensionamento e alinhamento de formas de maneira fácil e precisa; redimensionamento automático de páginas; alinhamento automático dos diagramas e dados a qualquer momento; e ajuste de layout.

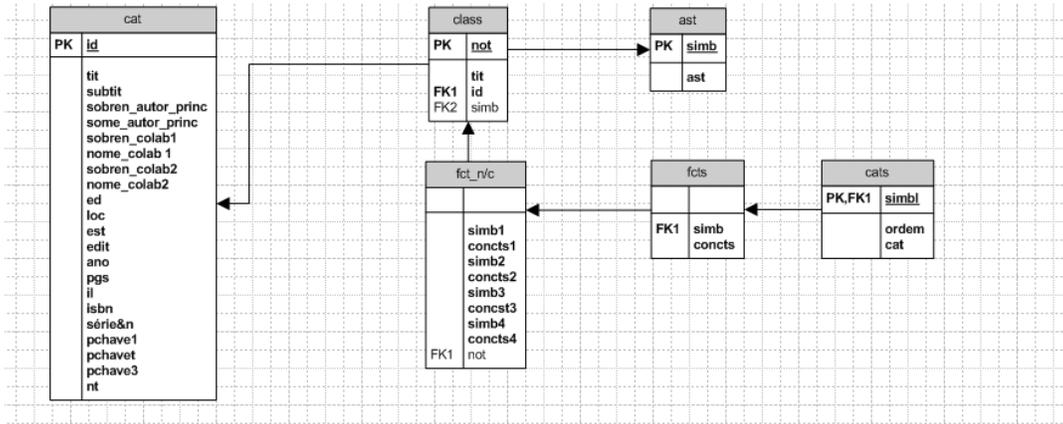


Figura 1: Diagrama de Modelo de BD.

Fonte: MS Visio 2007, Sistema Facetado (2011).

Conforme pode ser observado na figura 1, a proposta de implementação do *Sistema Facetado* está organizado em diagramas, que representam as respectivas tabelas (entidades) do sistema, que por sua vez apresentam os dados e seus atributos, e, por fim, seus relacionamentos.

A construção dos diagramas foi idealizada respeitando as características coletadas na modelagem conceitual, ou seja, conforme a atuação da TCF na formação dos assuntos, dentre os métodos *Dissecação (Dissection)*, o qual 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”; e *Desnudação (Desnudation)*, que também permitiu representar o núcleo específico de um assunto básico, ou de uma ideia isolada dos materiais consultados, sendo considerado o método mais importante para o sistema, pois a Instituição necessitava que a recuperação de informações fosse precisa, exigindo maior especificidade dos assuntos.

No caso da *Dissecação*, o diagrama teve que ser elaborado levando em consideração as características conceituais da “Educação Profissionalizante em Informática”, ou seja, inserir no sistema termos representativos ao citado assunto

básico. Por sua vez, na *Desnudação*, o sistema teria uma classificação ilimitada, permitindo que o usuário fosse o mais específico possível.

Apesar da primeira tabela, “Catalogação – cat”, ser a primeira do sistema e a que recebe as informações dos materiais, a tabela mais importante foi “Facetas_nível/ciclo – fct_n/c”, pois era o elemento responsável em apresentar as facetas e disponibiliza-las a *Classificação*, que por sua vez era relacionada ao *Cadastro*, por um lado, e ao *Assunto*, pelo outro. Por esse motivo a referida tabela teve que ficar no núcleo do BD, alimentando a multidimensionalidade nos momentos de classificação e recuperação da informação.

Na modelagem física parte da lógica, por meio da descrição das “estruturas físicas de armazenamento dos dados, tais como: tamanho dos campos, índices, tipo de preenchimento destes campos, nomenclaturas, etc, projetados de acordo com os requisitos de processamento [...]” (MACHADO; ABREU, 1996, p. 25).

Esta é a etapa final do desenvolvimento do BD, preparando o sistema para ser usado pelo usuário. Nesta etapa, o nível de abstração é mínimo, pois a estrutura de armazenamento do BD e seus dados serão apresentados de maneira definitiva, ou seja, será apresentada a versão definitiva ao usuário.

Selecionadas os dados que constituirão o sistema e sua organização lógica, a próxima etapa é escolher os recursos de *software*, conforme as características oferecidas pelo SGBD. Neste ponto, usamos os Sistemas Operacionais (SO) *MS Windows XP* e *MS Windows 7*. Este último foi uma escolha própria, por acompanharmos os recursos de *software* mais modernos. Já o *Windows XP* era o SO usado pela Instituição.

Quanto ao SGBD, selecionamos o *MS Access*, por desenvolver aplicações sob a abordagem relacional, ou seja, adota o tipo de BD relacional em sua modelagem lógica. Além disso, foi uma solicitação da Instituição, já que possui a assinatura digital do Pacote *Microsoft Office*, versões 2003 e 2007. A versão utilizada para o desenvolvimento do sistema foi a 2007 e testado na versão 2010, seguindo a mesma compreensão de atualidade tecnológica imposta ao SO. O sistema funcionará, na Instituição, tanto em computadores com a versão 2003 do Pacote *MS Office*, como também na versão 2007. Até o momento da implementação do *Sistema Facetado*, não havia uma padronização de recursos de *software* nos computadores da Instituição.

No que se refere às estruturas de programação, adotamos o *Structured Query Language* (SQL), para a realização das consultas, e o Visual Basic for Applications (VBA), para a programação mais especializada de partes do sistema, especialmente as relacionadas a classificação e a pesquisa.

O *MS Access* é um SGBD relacional e por isso sua programação ocorreu por meio de tabelas. No protótipo, conforme as definições da modelagem física, quanto aos recursos de *software*, foram criadas seis tabelas e relacionadas segundo suas associações e necessidades do sistema.

As tabelas criadas foram *Cat* para *Cadastro* (catalogação); *Classif* para *Classificação*; *Assuntos*, contendo todos os assuntos dos materiais presentes na Instituição que, no caso deste trabalho, iniciamos pelo assunto Informática; *Categorias*, que possui os símbolos e suas respectivas categorias (PMEST); *Facetas*, formada pelos símbolos das categorias e seus respectivos conceitos; e *Facetas_nível/ciclo*, tabela em que ocorrerá a classificação facetada conforme o material e necessidades da Instituição.

Além das seis tabelas, também foram criadas três consultas para realizarem as pesquisas no sistema: *Cons_Cat* (consultar o *cadastro*), para consultar informações do material após a pesquisa; *Cons_Classif* (consultar a *classificação*), necessária para associar o material à seus conceitos no momento da pesquisa; e *Cons_Rec* (módulo de recuperação da informação), considerado um dos objetos mais importantes do sistema, pois é a partir desta consulta que as informações serão recuperadas.

Também foram criados formulários para a interface do sistema e uso, como Menu, Categorias, Classificação, Recuperação etc., além de dois relatórios: um que listará todos os materiais cadastrados e outro que organizará os conceitos por categorias fundamentais.

Depois de criadas as tabelas, relacionamos conforme é apresentado na figura 2. Podemos observar a existência de uma sequência em que o sistema apresenta para que a informação seja recuperada pelo usuário.

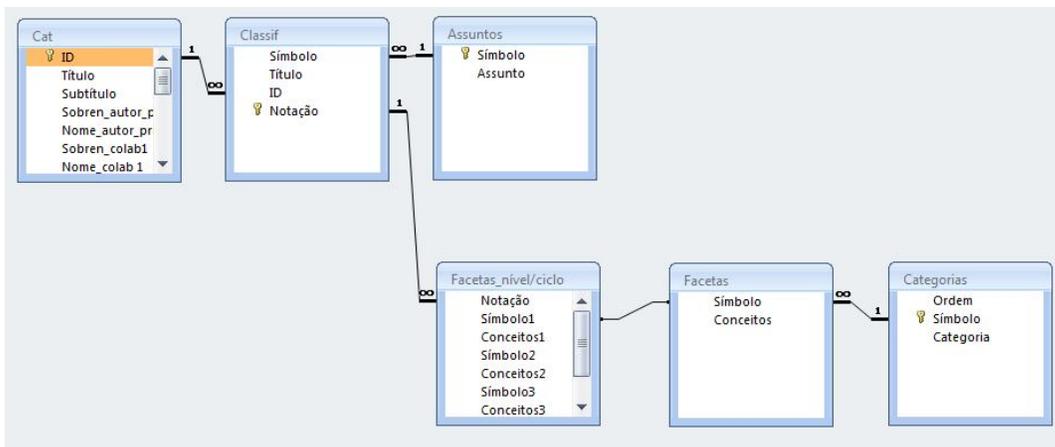


Figura 2: Relações entre as tabelas do Sistema Facetado.
 Fonte: MS Access 2007, Sistema Facetado (2011).

As tabelas e relacionamentos apresentados são uma tradução da modelagem lógica criada na etapa anterior. Primeiramente, o material é cadastrado no sistema. Em seguida, deve ser classificado em algum assunto e receber uma notação. Na tabela *classificação*, ocorrem três relações, sendo a primeira listar informações do material cadastrado, a segunda listar os assuntos presentes no sistema e a terceira adotar as facetas correspondentes ao assunto determinado.

Após o material receber uma notação, que tem apenas como função um controle de ordem do sistema e não a mesma atribuição presente nas Bibliotecas Físicas, ele deve ser classificado segundo a TCF, realizando a terceira relação. Para que o material receba a classificação facetada, é necessário selecionar conceitos, que por sua vez estão organizados em categorias.

Completado este ciclo, o material está classificado e pronto para ser recuperado pelo usuário. Com isso, podemos perceber a importância atribuída a classificação da informação para solucionar os problemas de recuperação da informação na Instituição.

O sistema apresenta uma tela inicial contendo um Menu (figura 3), organizado em três seções. A primeira seção chama-se *Módulos*, contendo *Cadastro* (catalogação), *Classificação* e *Pesquisa*. A segunda seção é formada pelos *Relatórios*, formado por *Materiais*, que listará todos os materiais cadastrados no sistema; e *Conceitos*, que apresentará uma lista de todos os conceitos presentes no sistema, organizados por categoria.

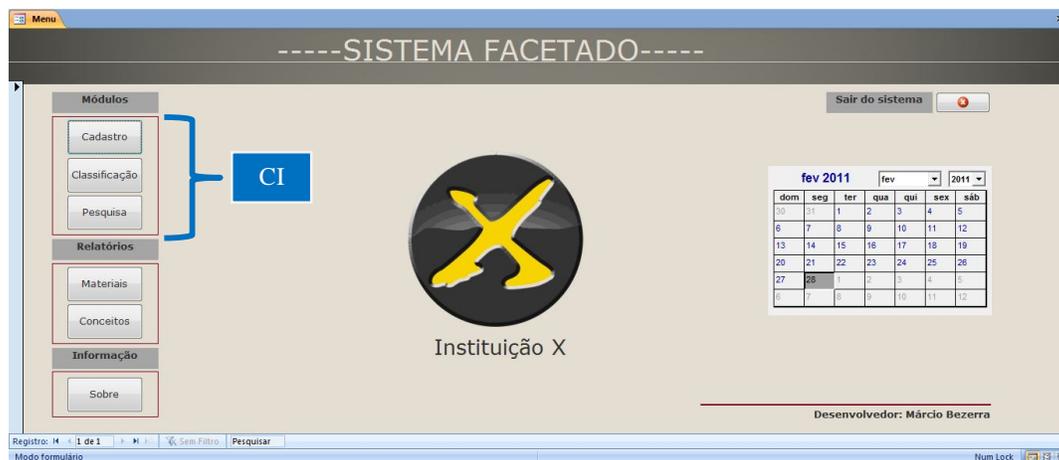


Figura 3: Menu Principal.

Fonte: MS Access 2007, Sistema Facetado (2011).

Além dos botões supracitados, estão presentes um botão para sair do sistema (X), navegação, ocorrendo também por guias ou abas. A cada módulo aberto, uma guia (aba) é afixada, ficando a escolha do usuário deixar o módulo aberto ou fechá-lo.

Na seção *Módulo*, o primeiro botão remete o usuário ao *Cadastro* de Materiais. Este módulo atua segundo a representação descritiva da informação, ou seja, a catalogação, conforme é apresentado na figura 4. Os campos para preenchimento dos dados são os mesmos apresentados pelo Código de Catalogação Anglo-Americano 2ª edição (AACR2), subsídio da CI para e elaboração deste módulo.

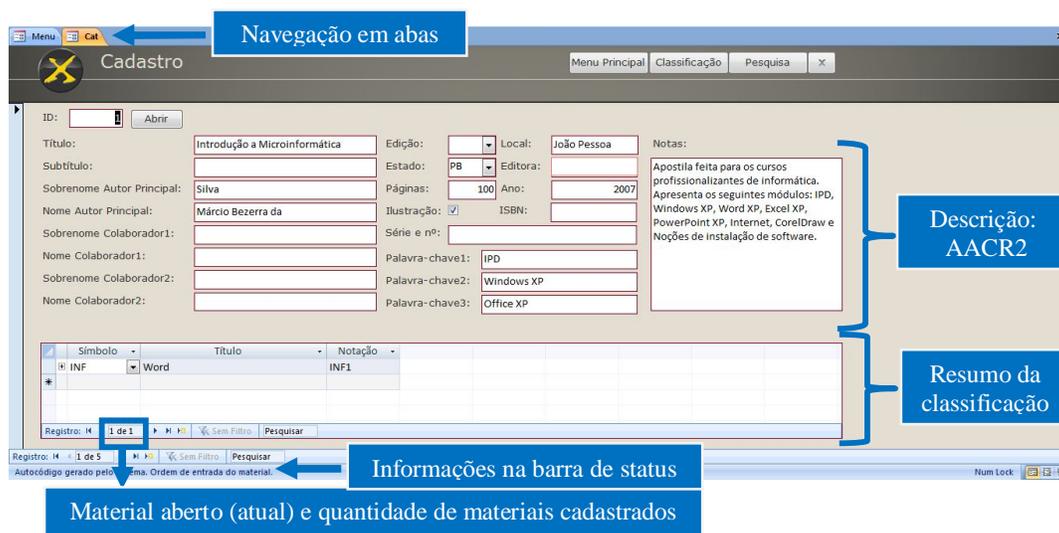


Figura 4: Módulo Cadastro de Materiais.

Fonte: MS Access 2007, Sistema Facetado (2011).

Na parte inferior da tela, figura 4, temos um resumo do material classificado e sua respectiva notação de sistema. Além disso, na barra de status, os usuários são informados sobre o preenchimento de cada campo por meio de mensagens.

Para que as informações da notação, de determinado material, constem no módulo *Cadastro* de Materiais, é necessário acessar o módulo *Classificação* (figura 5), que representa outros subsídios da CI: os Sistemas de Organização do Conhecimento (SOCs) e a classificação. Neste módulo, o usuário seleciona o símbolo do assunto e em seguida a ID² do material cadastrado no módulo anterior. No campo título, conforme a seleção da ID, automaticamente surge o título do material cadastrado. Possuindo o símbolo do assunto e o número da ID do material, é formada a sua notação (INF [Símbolo do Assunto] + 1 [ID] = INF1).

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

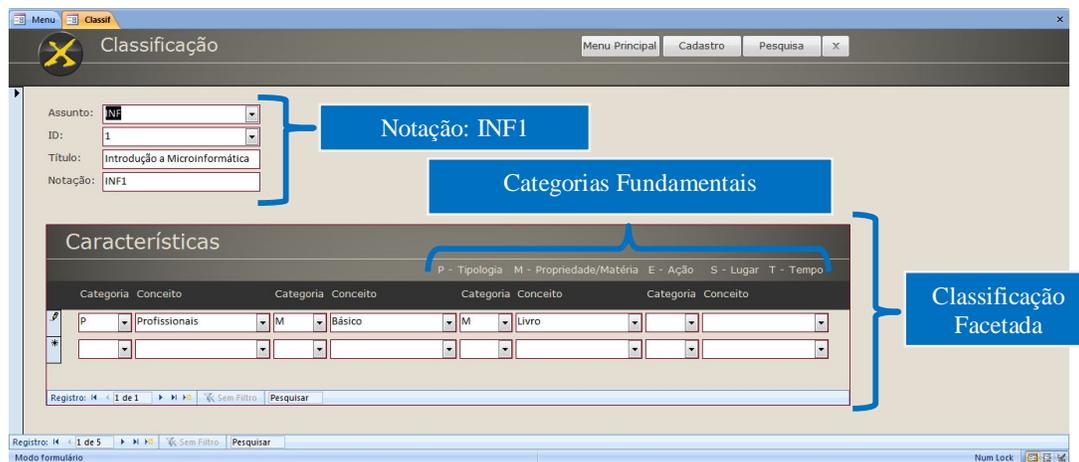


Figura 5: Módulo Classificação.

Fonte: MS Access 2007, Sistema Facetado (2011).

Na parte *Características*, o usuário determinará as facetas para cada material, a ser chamado de conceitos. Os conceitos estão organizados em categorias, que receberam outras denominações para representarem, de maneira mais fidedigna, a realidade do Setor Pedagógico. Esta atualização foi realizada na modelagem conceitual, a saber: “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” (DA SILVA; NEVES, 2011, p. 552). Conforme é apresentado na figura 6, verificamos que, ao selecionar de uma categoria pelo usuário, os respectivos conceitos surgem para realização da classificação facetada.

Depois de realizada uma classificação, o sistema remete o usuário para uma segunda seleção de categoria e assim sucessivamente, de maneira infinita, flexível e sem alterar a estrutura organizacional do sistema. Neste ponto, duas características da TCF foram inseridas de forma fidedigna ao *Sistema Facetado*. O uso do PMEST, em alguns momentos, pode apresentar um assunto que necessite da presença de uma mesma categoria mais de uma vez. Quando isso ocorre, temos os *Levels* (níveis) e os *Rounds* (*ciclos*). Os *níveis* ocorrem diante da aparição de duas ou mais categorias relacionadas ao mesmo assunto principal. Os níveis são indicados pela sigla de sua categoria, seguidos do algarismo arábico que indicará a ordem sequencial. Assim, na classificação “INF1[P],[M],[M2]” verificamos segundo nível da Matéria – Propriedade (M2), do título “Introdução a Microinformática” (INF1).

No caso dos *ciclos*, ocorrem quando uma categoria precisa reaparecer após o aparecimento de outro tipo de categoria. Para exemplificar, retomemos ao exemplo do título “Introdução a Microinformática” (INF1), acrescentando um clique após o segundo

nível da Matéria – Substância. Na classificação “INF1[P]:[M],[M2],[2P]” constatamos, após Matéria – Substância, o segundo ciclo da Personalidade – Tipologia (2P) do título “Introdução a Microinformática” (INF1).

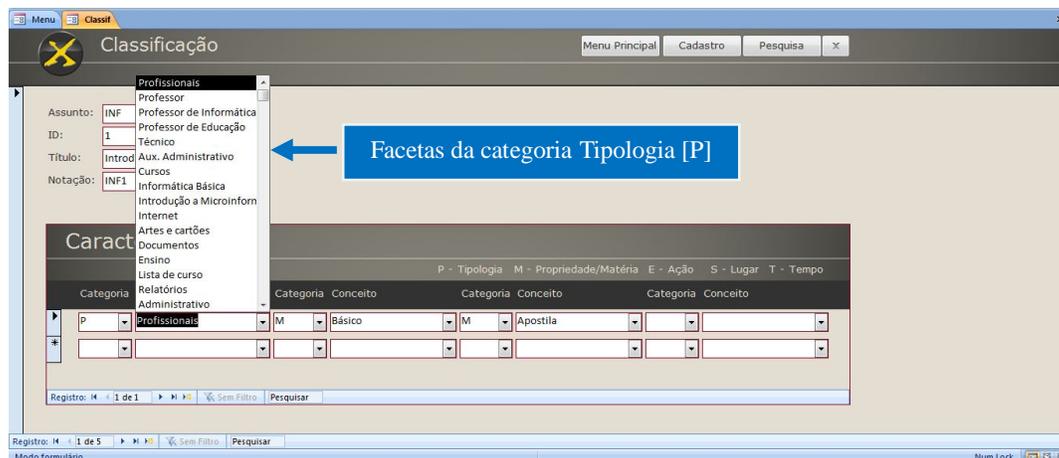


Figura 6: Características dos materiais segundo categorias.

Fonte: MS Access 2007, Sistema Facetado (2011).

Esta é uma ação fundamental para o protótipo, pois, conforme a classificação facetada realizada, os materiais serão recuperados de maneira mais precisa. Além disso, os usuários terão um “mapa categorial” contendo todas as categorias e conceitos para orientação no momento da classificação facetada. Em seguida, o material estará pronto para ser recuperado. Ambas as ações, *Cadastro* e *Classificação*, refletem no módulo *Pesquisa*. Ao buscarmos por *XP*, o sistema apresentou a ocorrência de 6 materiais, conforme apresentado na figura 7.

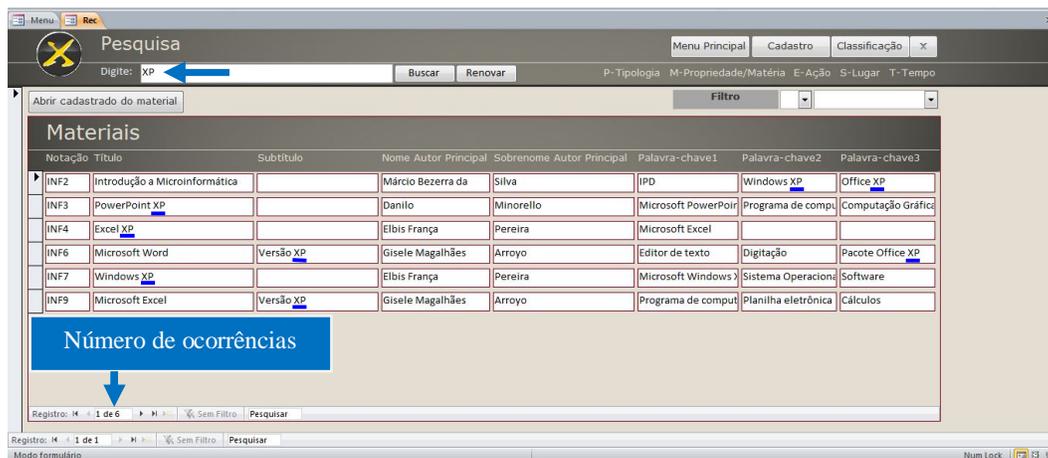


Figura 7: Módulo Pesquisa.

Fonte: MS Access 2007, Sistema Facetado (2011).

Sua construção se baseou na validação do protótipo conforme as necessidades dos usuários. Será a partir destes sujeitos que o sistema buscará as informações desejadas, ao invés de determinar que o sistema faça perguntas. Depois de realizada a busca, no item *Materiais* surgirá aqueles que atenderam ao que foi digitado pelo usuário. As informações do material são apresentadas de maneira resumida. Para tanto, foi criado um algoritmo, chamado *algoritmo_pesquisa*, que fizesse, em primeiro nível, uma busca pelos campos do módulo *Cadastro* considerados mais importantes pelos usuários do sistema: Notação, Título, Subtítulo, Sobrenome Autor Principal, Nome Autor Principal, Palavra-chave1, Palavra-chave2 e Palavra-chave3.

Caso o usuário tenha o desejo de filtrar a sua busca, diferentemente do que ocorrem nos SRIs mais conhecidos, neste módulo, o usuário poderá digitar mais uma vez, filtrando a ocorrência do resultado preliminar a partir de o *algoritmo_filtro_ocorrencia*, em segundo nível, cujo intitulamos de *Filtragem de Ocorrência*. Voltando a ilustração da figura 29, ao usarmos a *Filtragem de Ocorrência* pelo termo *Windows*, do total inicial de 5 ocorrência (*XP*), o *algoritmo_filtro_ocorrencia* reduzirá para 2 materiais que possuem o termo escolhido. Observe a figura 8:

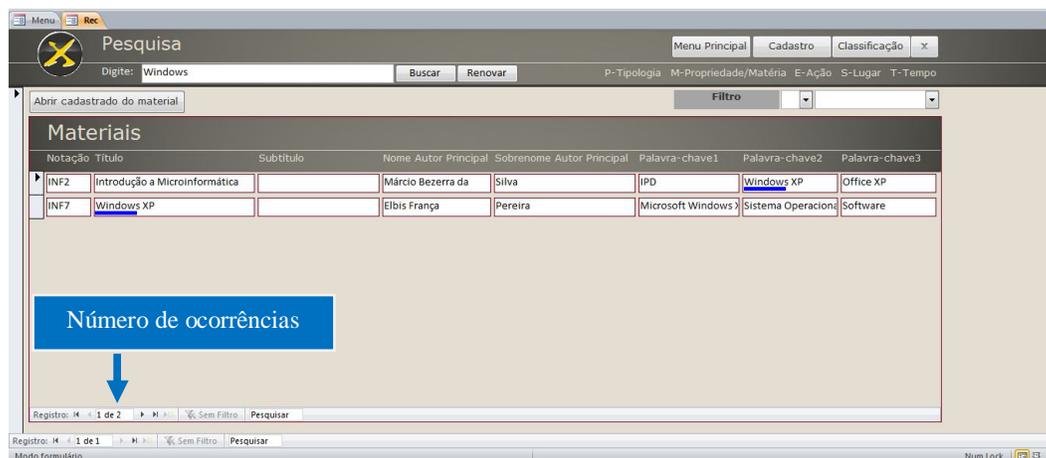


Figura 8: Módulo Pesquisa.

Fonte: MS Access 2007, Sistema Facetado (2011).

Além desta filtragem, encontramos na TCF uma grande aliada, pois, ao selecionar determinada categoria e, posteriormente, um conceito, o sistema filtrará o resultado da busca, tornando-o ainda mais específico. Neste sentido, criamos o algoritmo de terceiro nível *algoritmo_faceta*, que por sua vez é usado quantas vezes forem necessárias na busca, realizando uma *Filtragem de Ocorrência* dentre das facetadas. Neste caso, a busca

apresentará apenas 1 ocorrência, pois a filtragem ocorreu pela categoria Material – Matéria – Apostila, conforme é apresentado na figura 9. O outro material não poderia aparecer na ocorrência, pois, de acordo com a filtragem facetada, o mesmo foi classificado como um livro e não uma apostila.



Figura 9: Busca filtrada por categoria e conceito no Módulo Pesquisa.

Fonte: MS Access 2007, Sistema Facetado (2011).

Selecionado o material, seguindo orientações da Instituição, criamos um botão, chamado *Abrir Cadastro de Material* (figura 10) para que seja digitado o número da ID do respectivo material e, assim, ser direcionado ao módulo *Cadastro*. Após conferir se realmente o material selecionado atenderá suas necessidades, o usuário clicará no botão *Abrir*, caso esteja no formato digital, ou dirigir-se-á estante, onde o material encontra-se organizado. No caso da última busca, deve-se digitar a ID “2”, correspondente a notação INF2.

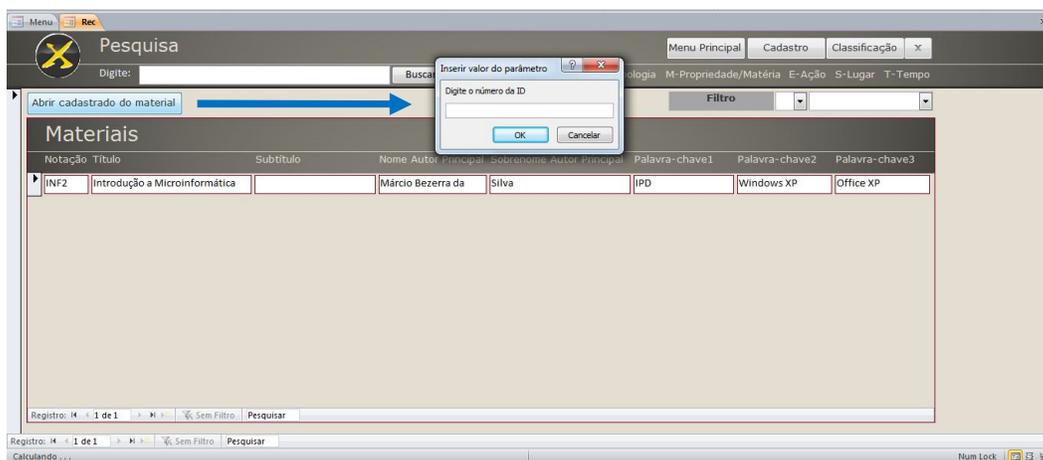


Figura 10: Botão para abrir o cadastro do material selecionado.

Fonte: MS Access 2007, Sistema Facetado (2011).

Na seção seguinte, são disponibilizados aos usuários dois relatórios. O primeiro é uma lista de todos os *Materiais Cadastrados* no sistema, conforme os dados selecionados pela Instituição. Também é apresentada a quantidade, data da solicitação e quantidade de páginas do relatório.

Já o segundo relatório é uma lista de todos os *Conceitos* presentes na classificação facetada do protótipo, organizados por categorias. Esta lista será de grande utilidade para o usuário responsável pela classificação dos materiais cadastrados no sistema. Para auxiliar o usuário, também está presente uma espécie de legenda, para auxiliá-lo na compreensão da leitura dos dados. Além disso, assim como no relatório de *Materiais Cadastrados*, também é apresentada a data da solicitação e a quantidade de páginas do relatório.

Por fim, a seção chamada de *Informação* foi uma solicitação da Instituição para alertar os usuários quanto aos objetivos do sistema e que os mesmos não solicitem materiais antes de realizar consultas, para verificar a presença ou não de determinado material no acervo do Setor Pedagógico.

A construção do *Sistema Facetado* foi inspirada por modelos de Sistemas de Automação em Biblioteca (SABs). Naturalmente, ocorre à associação a classificação de livros, tarefa esta especializada dos bibliotecários. No contexto apresentado, para Litton (1975, p. 1), “[...] devemos recordar que o homem está constantemente ordenando coisas e que a faculdade de classificar é um dos processos mentais essenciais”. Diante disso, assim como é defendido pelo autor, entendemos que a classificação facetada vai além de sua utilização em livros, ordenando e classificando os mais variados suportes informacionais, especialmente os digitais, ao qual, é um dos focos deste trabalho.

Caso seja de interesse ou necessidade específica, é possível criar um índice com todos os termos e suas respectivas notações (BARBOSA, 1972). Neste caso, foi criado um Mapa Categorical do *Sistema Facetado* para auxiliar os usuários na classificação dos materiais, assim como informamos em linhas anteriores.

4 CONSIDERAÇÕES FINAIS

Ao longo do presente trabalho, resolvemos discutir o que se encontrava por trás da interface do sistema, ou seja, a organização/classificação, para fins de recuperação. Para tanto, a união entre a Ciência da Informação (CI) e Ciência da Computação (CC), ambas interdisciplinares, nos ofereceram os elementos para a construção do sistema.

As contribuições oferecidas por ambas as áreas são inúmeras. No caso deste trabalho, buscamos adotar estudos da CI numa aplicação da CC, realizados com o objetivo de adotar formas de classificação em ambientes digitais, como foi o caso do *Sistema Facetado*, com o objetivo de catalogar os materiais do Setor Pedagógico da Instituição, classifica-los e, por fim, recupera-los. Nas práticas da CI, enraizadas na Biblioteconomia, adotamos a Teoria da Classificação Facetada (TCF), do matemático e indiano Shiyali Ramamrita Ranganathan (1892-1972), em um ambiente digital, neste caso, um BD. Após estudos, 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 que isto altere a estrutura do sistema.

Depois dos estudos e o desenvolvimento do *Sistema Facetado*, ao longo dos testes de pesquisa, constatamos que a análise de determinado assunto por facetas significou que o mesmo era visualizado pelas manifestações das características ou facetas atribuídas a ele, assunto, o que, no caso deste trabalho, chamamos de conceitos. Estas manifestações são muito importantes, pois o *Sistema Facetado* tornou-se multidimensional e ilimitado, representados pelas facetas e organizadas por categorias.

Outro ponto forte foi na etapa de recuperação da informação. Além da forma tradicional de busca por digitação, acrescentamos uma *Filtragem de Ocorrência* dentro da própria busca (preliminar), permitindo que os usuários refinassem a sua busca digitando outros termos. Vendo os benefícios desta técnica, adotamos também uma filtragem por facetas. Neste sentido, o usuário escolheria uma categoria fundamental e uma de suas facetas para realizar a filtragem, quantas vezes fossem necessárias. Com isso a ideia de Ranganathan de apresentar características aos assuntos, permitindo benefícios à recuperação, foi concretizada. De fato, a TCF permite tornar o assunto mais específico e, com isso, uma busca mais precisa.

A grande questão deste trabalho encontra-se na ausência de modelos e/ou estudos empíricos que apresentassem metodologias e testes. O desenvolvimento do protótipo de BD partiu na imaginação e conhecimento profissional em BD. A partir das leituras e virtualização (mental) de como a TCF poderia ser inserida no sistema, escolhemos a programação em BD Relacional e foi justamente este método que nos permitiu a concretização do presente trabalho, apresentado tecnicamente neste estudo.

Apesar dos resultados satisfatórios, muitos foram os percalços enfrentados. A Instituição possui uma administração pública, afetando o cronograma de pesquisa não pôde ser atendido. Situações como agenda política e eleições nos fizeram reorganizar as atividades de pesquisa diversas vezes. Neste período, três BDs foram criados, pois passamos por três administradores diferentes, divergindo de ideias. Além disso, desde o início acordou-se que a Instituição deveria dar o crivo para o andamento da pesquisa.

Por fim, o *Sistema Facetado*, no final do ano de 2011, recebeu uma atualização e os usuários começaram a inserir materiais de outros cursos. Neste período, pelo sucesso da proposta, a Instituição adquiriu os seus direitos legais. A partir destes acontecimentos, esperamos novos estudos, novas propostas, com fins a melhorias nas formas de recuperação dos ambientes digitais. Sendo assim, após a leitura do presente estudo, esperamos que esteja mais clara a importância da organização para o sucesso de qualquer sistema.

Ao continuarmos na árdua caminhada ocorrida para a realização deste trabalho, almejamos ampliar esta discussão para ambientes mais gerais, abrangentes, como é o caso da Web.

Abstract: Presents the insertion of the Ranganathan's Theory of Faceted Classification (TCF) in prototyping Database (DB) from the data modeling, specifically in the programming steps: logical and physical. Discusses modeling of data on BD as a technique that should represent the reality of the environment that the system will be inserted. The object of a methodological path formed by the Research Applied, Exploration and Bibliographical, and qualitative approach, to present the application of faceted classification of knowledge organization, focusing on the recovery of information in the database, but also investigate the possibility of application of faceted classification with data modeling in digital environments and expose an overview of using the prototype BD. Collection, as search results, the contributions of TCF in the logical and physical modeling, respecting the methods of forming the subject of Ranganathan, Dissection and Desnudation, done in conceptual modeling (other work). Makes in physical modeling the choice of software resources, the manufacture of the modules in the system Registry (cataloging), Classification and Search, and contributions of TCF in the use of unlimited facets of the occurrences and filtering search, making searches more exacts. It is concluded that subsidies for the development of the Faceted System was TCF and modeling data in BD, even with satisfactory results, the big question lies in the absence of models and/or empirical studies that present methods and tests, giving us as a starting point the imagination and professional knowledge in BD. In the end longing desire of other studies, other models, which can help us extend this discussion to more general environments, comprehensive, such as Web.

Keywords: Ranganathan's Theory of Faceted Classification. Data Modeling. Logic Modeling. Physical Modeling.

REFERÊNCIAS

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 12 jul. 2012.

BATTIST, J. C. F. **Treinamento: Access** – Técnicas Avançadas. Santa Maria, 2004. (CD-ROM).

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 jul. 2012.

CAMPOS, M. L. A. **A organização de unidades do conhecimento em hipertextos**: 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.

DAHLBERG, I. Teoria da classificação, ontem e hoje. In: Conferência Brasileira de Classificação Bibliográfica, Rio de Janeiro, 12-17 de setembro de 1972. **Anais**. Brasília, IBICT/ABDF, 1979. v. 1, p. 352-370. Disponível em: <http://www.conexaorio.com/bit/dahlbergteoria/dahlberg_teoriam.htm>. Acesso em: 12 jul. 2012.

DATE, C. J. **Introdução a Sistemas de Bancos de Dados**. Rio de Janeiro: Elsevier, 2003. Tradução da 8ª edição americana.

ELMASRI R., NAVATHE S. B. **Fundamentals of database systems**. 2. ed. The Benjamim/Cummings, 1994.

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

HEUSER, C. A. **Projeto de Banco de Dados**. 5. ed. UFRGS: Editora Sagra Luzzanato, 2004. (Série Livros Didáticos).

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

KORTH, H. F., SILBERSCHATZ, A. **Sistema de Banco de Dados**. 2 ed. São Paulo: Makron Books, 1995.

KUMAR, K. **Theory of classification**. 2. ed. New Delhi: Vikas Publishing House, 1981.

LANGRIDGE, D. **Classificação**: uma abordagem para estudantes de Biblioteconomia. Tradução de Rosali Pacheco Fernandez. Rio de Janeiro: Interciência, 1977.

LIMA, G. A. B. **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.

LITTON, G. **Classificação e catalogação.** São Paulo: McGraw-Hill, 1975.

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

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

RANGANATHAN, S. R. **Colon Classification.** Bombay: Ásia Publishing House, 1963. 126p.

_____. **Heading and canons: comparative study of five catalogue codes.** Madras, S. Viswanthan, 1955. 300 p.

_____. **Prolegomena to library classification.** Bombay: Asia Publ. House, 1967.

RICARTE, I. L. M. **Programação Orientada a Objetos com C ++.** Versão preliminar (apostila). Campinas: Universidade Estadual de Campinas, 2001.

ROBREDO, J. **Da ciência da informação revisitada aos sistemas humanos de informação.** Brasília: Thesaurus, 2003.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, v. 1, n. 1, p.41-62, 1996. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>>. Acesso em: 12 jul. 2012.

DA SILVA, M. B. A aplicação da folksonomia em sistema de informação. In: IV Encontro Brasileiro de Arquitetura da Informação – EBAI, São Paulo, 12-13 de novembro, 2010. **Anais.** Disponível em: <http://www.congressoebai.org/wp-content/uploads/ebai10/EBAI10_artigo10.pdf>. Acesso em: 12 jul. 2010.

_____; NEVES, Dulce A. de Brito A aplicação da teoria da classificação facetada em banco de dados, através da modelagem conceitual. XII Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), Brasília (DF), 23-26 de outubro de 2011.

Anais eletrônicos. Brasília: Thesaurus, 2011. 540-557. 1 CD-ROM.

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

YONG, C. S. **Banco de Dados – Organização, sistemas e administração.** São Paulo: Editora Atlas S. A, 1986.