



TAXONOMIA FACETADA NAVEGACIONAL: AGREGANDO VALOR ÀS INFORMAÇÕES DISPONIBILIZADAS EM BIBLIOTECAS DIGITAIS DE TESES E DISSERTAÇÕES

Benildes Coura Moreira dos Santos Maculan, Gercina Angela Borém de Oliveira Lima

Resumo: Este artigo apresenta o resultado de uma pesquisa de mestrado que investigou o uso de uma taxonomia facetada navegacional (TAFNAVEGA) como instrumento de organização da informação, para dar consistência e padronização aos conteúdos informacionais de teses e dissertações. Teve como objetivo facilitar o acesso às informações, a partir da navegação através dos termos da estrutura facetada. O ambiente de pesquisa foi uma biblioteca digital de teses e dissertações (BDTD) de uma unidade informacional específica, tendo como insumo as teses e dissertações particularmente oriundas da linha de pesquisa Organização e Uso da Informação (OUI). A TAFNAVEGA é composta por dez categorias fundamentais temáticas (CAFTE) e por um algoritmo para extração de conceitos, denominada matriz categorial temática para trabalhos acadêmicos. A metodologia incluiu o uso da teoria da análise de domínio para a identificação do domínio, do usuário da BD TD e de suas necessidades informacionais, assim como do Modelo de Leitura Documentária, de Fujita e Rubi (2006), para o desenvolvimento da matriz categorial para trabalhos acadêmicos. Para a criação das categorias fundamentais temáticas foi utilizado o método da análise de conteúdo, com a técnica da análise categorial temática. Os resultados demonstraram que o mecanismo TAFNAVEGA pode facilitar a exploração, busca e recuperação dos conteúdos dos documentos, dando acesso a dados tais como teorias, métodos e instrumentos de coleta de dados. Concluiu-se que esse instrumento auxilia o pesquisador em sua atividade profissional e permite maior visibilidade ao conteúdo disponível na biblioteca digital de teses e dissertações, sem sobrecarregá-lo de informações.

Palavras-chave: Taxonomia facetada navegacional. Navegação facetada. Busca e recuperação da informação. Biblioteca de teses e dissertações. Matriz categorial para trabalhos acadêmicos.

1 INTRODUÇÃO

O ser humano é essencialmente social, ligado a diferentes grupos de interesses, tradicionalmente transmitindo seu conhecimento a outro e também recebendo saberes de outros. A disciplina Ciência da Informação (CI) notadamente se ocupa em entender esses fluxos de informação, em distintos ambientes, desenvolvendo formas de facilitar seu acesso e uso (BORKO, 1968), o que permite expandir o conhecimento da humanidade acerca da ciência e tecnologia.



O saber científico é dinâmico e primordial no desenvolvimento da sociedade, num ciclo aberto de reconstrução, se transformando, aperfeiçoando e ampliando. A ciência passa sempre por revisões em suas teorias, seja por não mais se adaptarem à realidade empírica existente ou devido à descoberta de novos fenômenos. Sendo assim, as teorias científicas são falíveis e estão, continuamente, sendo substituídas ou recondicionadas, principalmente para se ajustarem à época histórica ou científica. A evolução da ciência acontece com a possibilidade de replicação, quando um novo conhecimento emerge a partir de algo já descoberto. Nesse sentido, a comunicação científica tem importante papel, pois permite o acesso às diferentes abordagens teóricas, epistemológicas e práticas que estão sendo utilizadas nas pesquisas, e também aos resultados obtidos nas mesmas.

Na academia, há manifesta preocupação com a comunicação e difusão do conhecimento advindo de pesquisas já concluídas, como forma de obter visibilidade institucional, legitimando e promovendo maior prestígio entre os pares. Nesse sentido, uma significante iniciativa de instituições de ensino superior é a criação de bibliotecas digitais para a disponibilização de teses e dissertações. Com isso, pretende-se facilitar o acesso a esse tipo de documento e evitar esforços replicados em pesquisas sobre mesmo tema ou temas semelhantes.

Contudo, é importante destacar que ainda pode ser considerado difícil recuperar os conteúdos detalhados das pesquisas finalizadas. As teses e dissertações possuem conteúdo especializado, utilizando linguagem formal e específica para a área ou domínio, através de uma terminologia própria. Não obstante, são produzidas em linguagem natural, o que ocasiona casos de sinonímia (palavras com significados iguais) ou até mesmo de homonímia (identidade fonética e/ou gráfica de palavras com significados diferentes). Essa particularidade causa dificuldade na comunicação e compartilhamento do conhecimento registrado nesses documentos. Esse fato é comumente agravado por problemas na indexação dos documentos, pois as palavras-chave atribuídas nem sempre são representativas de toda a temática tratada ou os resumos informativos estão incompletos, ocorrências essas já amplamente ressaltadas na literatura da área (BORKO; CHATMAN, 1963; VIEIRA, 1988; NOGUEIRA, 1997; GIL LEIVA; FUJITA; RUBI, 2008; ELIEL, 2008; SOUZA, 2009).

Outra dificuldade está relacionada à busca avançada, geralmente realizada através do uso da lógica *booleana* (*or*, *and* e *not*), utilizada para ampliação ou refinamento de resultados. Esse recurso pode ajudar na recuperação de informações mais específicas em uma biblioteca digital, mas a maior parte dos usuários finais não está apta a utilizar tal opção. Os motivos para esse impedimento podem advir de diferentes fatores, seja porque o usuário não consegue acompanhar as mudanças tecnológicas (barreira tecnológica), por comodismo (barreira psicológica) (FIGUEIREDO, 1987), ou mesmo porque o utente considera as buscas *booleanas* de difícil compreensão (BAEZA-YATES; RIBEIRO-NETO, 1999).

A decorrência desse cenário justificou a pesquisa de mestrado intitulada “Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos”, que criou uma taxonomia facetada para navegação, como opção de busca em bibliotecas digitais de teses e



dissertações. O recurso prioriza as necessidades de informação dos pesquisadores, com vistas a promover acesso mais facilitado aos resultados detalhados das pesquisas e dar maior visibilidade ao conhecimento científico acumulado nesse tipo de documento.

Neste artigo, primeiramente são apresentadas as características textuais de documentos acadêmicos do tipo teses e dissertações. Em seguida, descreve-se o desenvolvimento da taxonomia facetada navegacional (TAFNAVEGA), com sua conceituação, princípios de navegação hipertextual e os procedimentos da metodologia utilizada. Por fim, demonstram-se os resultados e conclusões.

2 DOCUMENTOS ACADÊMICOS DO TIPO TESES E DISSERTAÇÕES

Para a CI, o documento é a materialidade de um registro informacional que proporciona identificar e apreender seu conteúdo (RABELLO, 2009), pois um “documento” pode ser, desde a concepção de Otlet, qualquer objeto de informação que tenha caráter representativo (CUNHA; CAVALCANTI, 2008). Esse significado de “documento” está também expresso na norma que regula a prática da análise de documentos, a NBR-12676:1992. Essa norma determina que documento é “qualquer unidade, impressa ou não, passível de catalogação ou indexação”, e ainda acrescenta que:

Esta definição se refere não apenas a materiais escritos ou impressos em papel ou suas versões em microforma (p. ex.: livros, jornais, diagramas, mapas), mas também a suportes não-impressos (p. ex.: registros legíveis por máquinas, filmes, gravações sonoras), objetos tridimensionais e *realia* usadas como espécimens (NBR-12676, 1992, p.2).

Aliado a isso, Buckland (1991) indica a tangibilidade da unidade informacional (informação), ao apontar que ela pode ser disseminada em SRIs através de três distintos usos: 1) informação como processo (transmissão e comunicação); 2) informação enquanto conhecimento (reduzir incertezas através do uso); 3) informação na perspectiva de objeto (qualquer “coisa”).

Os documentos acadêmicos compreendem todo tipo de produção da comunidade científica. Incluem-se nesse arrolamento os documentos abordados neste artigo, do tipo teses e dissertações, que têm como característica apresentar o resultado de pesquisa com tema único e escopo bem delimitado. A diferença entre esses dois tipos de documentos é pouco significativa. As teses devem oferecer uma contribuição real para a questão levantada, com alto grau de profundidade de estudo, enquanto as dissertações devem priorizar a demonstração da capacidade de aplicação de método científico e a habilidade de sistematização do aluno, sem que seja esperado dele um aporte original. Esses documentos têm caráter didático (MARCONI; LAKATOS, 2010) e devem respeitar as regras e normas técnicas para a normalização textual, de modo a facilitar a comunicação e o compartilhamento das informações contidas neles.

No Brasil, a Associação Brasileira de Normas Técnicas (ABNT), é o órgão responsável por essa regulamentação técnica, sendo ainda representante de entidades internacionais tais como a International Organization for Standardization (ISO) e International Electrotechnical Commission (IEC). No tocante à elaboração dos documentos acadêmicos do tipo teses e dissertações, as diretrizes que prevalecem são as da Norma NBR 14724:2011. Outras normas complementam-na, quais



sejam: NBR 10520:2002 (Citações), NBR 6023:2002 (Informação e documentação – Referência – Elaboração), NBR 6024:2003 (Numeração progressiva das seções de um documento escrito), NBR 6027:2003 (Sumário), NBR 6028:2003 (Resumo), NBR 6034:2004 (Índice), NBR 12225:2004 (Títulos de lombada).

Segundo a Norma ABNT NBR 14724 (2011, p.8), as teses e dissertações possuem três partes fundamentais: parte introdutória, desenvolvimento e uma conclusão. A segunda parte, de desenvolvimento da pesquisa, é a mais complexa do documento textual e inclui quatro elementos principais: 1) revisão de literatura (outros estudos acerca do tema estudado); 2) descrição detalhada da metodologia (pressupostos, amostra, métodos, técnicas, instrumentos e processos); 3) resultados obtidos durante o desenvolvimento da pesquisa; 4) discussão dos resultados alcançados. Nesse contexto, o desenvolvimento da taxonomia facetada navegacional (TAFNAVEGA) teve como parâmetro essa estrutura textual dos documentos acadêmicos do tipo teses e dissertações.

3 DESENVOLVIMENTO DA TAXONOMIA FACETADA NAVEGACIONAL

A TAFNAVEGA foi desenvolvida a partir do universo de pesquisa representado pelas teses e dissertações defendidas entre os anos 1998 e 2009, no Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Escola de Ciência da Informação (ECI), da Universidade Federal de Minas Gerais (UFMG). No arrolamento desses documentos obteve-se um total de 290 trabalhos (62 teses e 228 dissertações).

Desse total, determinou-se o *corpus* de estudo, por amostragem não-probabilística (ou subjetiva), do tipo intencional (por julgamento), indicado para investigações de cunho exploratório, como no caso da pesquisa realizada, principalmente se o universo é homogêneo (MALHOTRA, 2001, p.316). Como resultado, obteve-se um *corpus* constituído por documentos provenientes da linha de pesquisa “Organização e Uso da Informação” (OUI), disponibilizadas no banco de dados da biblioteca digital de teses e dissertações da instituição (BDTD/ECI/UFMG), totalizando 41 documentos (12 teses e 29 dissertações).

3.1 Uma conceituação para taxonomia facetada navegacional

Em sua expressão mais tradicional, uma taxonomia apresenta um domínio na forma de uma simples hierarquia. Uma taxonomia facetada, por sua vez, apresenta um domínio em facetas e, em cada faceta, há uma hierarquia. Essa característica permite estabelecer relações entre as facetas, indicando a multidimensionalidade de um termo nesse mesmo domínio.

Em ambiente digital, o termo “navegação” refere-se à possibilidade interativa utilizada pelo usuário para acessar informações em um banco de dados. Isso pode ser atestado pela descrição do verbete “navegação” no Dicionário de Biblioteconomia e Arquivologia:

pesquisar (ou navegar) em um programa, procurar comandos, percorrer um documento e buscar informações [...] [que é o] processo dos usuários interagindo com um sítio visando satisfazer suas necessidades de informação [...], [navegando] em sítios por meio de busca e folheio de objetos com conteúdo (CUNHA; CAVALCANTI, 2008, p.257).



Percebe-se, nessa definição, a possibilidade de o usuário explorar o conteúdo de um banco de dados através da navegação. Bonilla (2005, p.139) acrescenta que “a navegação está baseada nas indexações e associações de idéias [sic] e conceitos, organizados sob a forma de *links*, os quais agem como portas virtuais que abrem caminhos para outras informações”. Esse sentido de “associação de ideias” sugere a imagem de uma navegação facetada, abordagem empregada para facilitar a interrelação e a recuperação de informações, que vem sendo denominada de “paradigma de navegação facetada” (*faceted browsing paradigm*) (BROUGTHON, 2006; TZITZIKAS; ANALYTI, 2007; SACCO; TZITZIKAS, 2009).

Tunkelang (2009) afirma que uma navegação facetada possibilita que o usuário elabore sua busca progressivamente e garante que ele a refine, apresentando as diferentes opções de escolhas e eliminando a possibilidade de “resultados vazios”, uma vez que somente mostra as facetas nas quais há conteúdos de informação.

Para a criação de uma taxonomia para navegação (navegacional), decompõe-se o conteúdo de um documento ou domínio em categorias e facetas, com base em conceitos e nas relações existentes entre as facetas, a partir de uma lógica de organização pré-estabelecida, tal como a página de busca da *Microsoft Network* (<<http://search.msn.com>>) ou da *HomeAdvisor* (<<http://homeadvisor.msn.com>>) (CONWAY; SLIGAR, 2002, *on-line*).

Partindo do exposto, pode-se definir “taxonomia facetada navegacional” como um sistema de categorias, onde cada categoria possui uma hierarquia de facetas e subfacetas, cuja subdivisão obedece a um mesmo critério pré-estabelecido, permitindo atribuir diferentes dimensões (multidimensional) a um objeto. Esse sistema é organizado sob a forma de *links*, o que proporciona recurso de navegação e acesso, possibilitando que o usuário interaja e explore o conteúdo informacional do banco de dados, filtrando e restringindo sua busca em conformidade com sua escolha de facetas (MACULAN, 2011).

3.2 Princípios de navegação hipertextual

O hipertexto (ou hipermídia) é um sistema para visualização de informações que pode conter itens tais como textos, áudio, vídeos ou imagens, ligados entre si de forma unidirecional ou multidirecional (*links*), de tal maneira que facilita a busca de informação.

O termo “hipertexto” foi cunhado em 1965, por Theodor Nelson, com o Projeto Xanadu (NELSON, 1999). Entretanto, considera-se o *ON-Line System* (NLS) como o primeiro sistema operacional de hipertexto, apresentado ao público em 1968, criado para facilitar o processo de armazenamento e recuperação de documentos em bibliotecas digitais.

Conklin (1988, p.424) conceitua hipertexto como “uma janela na tela, associada com objetos em um banco de dados, e ligações [*links*] estabelecidas com esses objetos, graficamente (rótulos) ou em banco de dados (metadado)”. Para ser considerado hipertextual, um sistema deve permitir aos usuários navegar por nodos de informação, explorando e obtendo as informações disponibilizadas em meio digital.



Segundo Furtado (2010), atualmente há duas vertentes sobre o conceito de hipertexto: uma de associação de ideias e outra relacionada à conexão de ideias. A primeira mencionada considera o hipertexto como uma associação não linear de ideias, quase casual, tal como pensava Bush. A outra vertente sugere que as conexões de um hipertexto são associações controladas, condicionadas à intencionalidade da comunicação de ideias de quem o cria, geralmente visando atingir não um único usuário, mas a uma comunidade de usuários. Nessa segunda vertente, subentende-se que a liberdade de navegação do usuário de um hipertexto está limitada às associações permitidas pelo sistema e imputadas pela intenção de comunicação do *designer*.

Adotando essa última premissa, conclui-se que em todo hipertexto há um propósito comunicativo pré-estabelecido, ainda que sua estrutura não sequencial e não linear permita escolher distintos caminhos e em diferentes ordens. Esse é o mesmo sentido de uma navegação facetada, na qual um usuário pode iniciar sua busca progressiva em qualquer ponto da taxonomia, dentre um conjunto limitado de itens e conexões. Organizado dessa forma, o banco de dados adquire a característica de um sistema hipertextual que facilita o acesso e a recuperação de informações.

A estruturação de informações através de taxonomias facetadas vem sendo comparada com interfaces de busca em *sites* comerciais (*e-commerce*). Cabe ressaltar, entretanto, que do ponto de vista teórico da CI, o recurso disponibilizado nessas interfaces nem sempre pode ser denominado de navegação facetada. Para a CI, uma taxonomia facetada é uma cadeia de relações, na qual se perde a noção de ordem hierárquica. Nessa cadeia não se pode determinar a maior importância de uma faceta em relação a outra, uma vez que elas estão associadas sob diferentes tipos de relações, sugerindo o sentido de cadeia de relações multidimensional e multidirecional (MACULAN, 2011). O que se percebe em distintas interfaces de buscas em sites comerciais são informações (ou produtos) estruturadas a partir de um único aspecto (atributo), ou sob até no máximo três aspectos. Considera-se que essa característica impede ou restringe demasiadamente a visão multidimensional e multidirecional do objeto em questão, condição obrigatória para que uma navegação possa ser considerada facetada, dentro dos princípios das teorias da CI.

Como exemplo de navegação facetada pode ser citado o sistema FLAMENCO (disponível em <<http://flamenco.berkeley.edu/>>), que destaca-se como um dos primeiros e o mais citado exemplo na literatura da área. Ele é composto por um conjunto de textos representados em um banco relacional, a partir de um modelo facetado pré-estabelecido, permitindo definir termos e criar hierarquias entre eles.

No desenvolvimento da taxonomia facetada navegacional (TAFNAVEGA) houve a preocupação em aplicar uma metodologia que mantivesse esses princípios – multidimensional e multidirecional –, buscando preservar as associações já existentes entre as diferentes partes textuais dos documentos do tipo teses e dissertações.



3.3 Metodologia de desenvolvimento da TAFNAVEGA

A TAFNAVEGA é constituída por dez categorias fundamentais temáticas (CAFTE) e pelo algoritmo representado pela Matriz Categorial Temática para trabalhos acadêmicos (matriz), utilizada para a análise conceitual nas teses e dissertações durante o “processo de extração de conceitos”. De forma sintetizada, os procedimentos para seu desenvolvimento estão apresentados a seguir.

3.3.1 Identificando o domínio, o usuário e suas necessidades informacionais

Como primeiro procedimento para o desenvolvimento da TAFNAVEGA, identificou-se o domínio e o usuário da BDTD, assim como as necessidades informacionais desse utente, aplicando-se o método da teoria da análise de domínio (TAD). Esse método foi principalmente baseado nos estudos de Hjorland e na literatura especializada sobre a comunidade científica, consistindo na realização de quatro etapas: 1) identificação do domínio; 2) análise de dados; 3) modelagem do domínio; 4) coleta e seleção de dados e validação.

A primeira etapa, de identificação do domínio, foi concretizada a partir das seguintes indagações: qual é o domínio? o domínio é conhecido? que tipo de informações esse domínio produz?. Para responder a essas questões houve a análise dos documentos para identificar a qual domínio ele pertence, qual sua função dentro desse ambiente e as razões subjacentes para os fluxos de informação nesse campo (HJORLAND, 2002). Ao término das análises nessa etapa, concluiu-se que o domínio é o programa de pós-graduação *stricto sensu* da ECI-UFMG, especificamente aquele que produz documentos acadêmicos do tipo teses e dissertações.

Em seguida, a etapa de análise de dados foi realizada para identificar os pontos centrais que caracterizam o domínio, estabelecendo relações e funções entre as informações, a partir dos seguintes questionamentos: quem produz as informações identificadas na etapa anterior? com qual propósito a informação é produzida? a quem interessa o uso das informações produzidas?. Para obter respostas a essas indagações, a análise foi baseada nas prescrições da teoria da análise de domínio (TAD) que afirmam que o fluxo de informação é determinado no seio da comunidade de discurso, a partir da observação de quem a produz, comunica e usa. A partir disso, verificou-se que o produtor de teses e dissertações é sempre um pesquisador de carreira ou de circunstância e que, como cientista, é membro da comunidade científica. A finalidade dessa produção é ser aprovado perante uma banca de doutores (área acadêmica), dentro de um mesmo campo de conhecimento ou entre campos interdisciplinares. Considerou-se que, em sentido *stricto sensu*, todo pesquisador aplica procedimentos metodológicos sistemáticos, que é parte inerente de sua atividade profissional e é exigido que esteja atualizado nos temas de seu campo de atuação científica, como forma de obtenção de reconhecimento por seus pares.

Na terceira etapa, de modelagem do domínio, efetivou-se a seleção dos tipos de informações mais adequados ao objetivo proposto, em conformidade com os dados identificados na etapa anterior.



Nessa seleção determinou-se que em atividade profissional o pesquisador necessita de informações sobre ideias, abordagens, teorias, métodos e técnicas que fundamentem e auxiliem em sua própria produção. Assim, seguindo a lógica da TAD, definiu-se: a) quem produz: a comunidade científica; b) quem comunica: as instituições (quando disponibilizam digitalmente); c) quem usa: a própria comunidade acadêmica, representada pelo pesquisador, em atividade profissional. Deliberou-se, ainda, que a Biblioteca de Teses e Dissertações (BDTD) em questão é disponibilizada pela instituição ECI-UFMG e o serviço de informação foi desenhado para atender, prioritariamente, aos pesquisadores.

Por último, na etapa de coleta e seleção de dados e validação, houve a validação da etapa anterior, utilizando a literatura especializada da área e o histórico do domínio, assim como consultas aos especialistas. A pergunta norteadora foi: quem é este “não-usuário total” a quem o conteúdo dos trabalhos acadêmicos, do tipo teses e dissertações, pode interessar?. O termo “não-usuário total” nesse contexto tem um sentido específico, que é o daquele usuário que não usa um serviço porque acredita que tem outros meios de conseguir a informação ou porque não está ciente de que há aquele tipo de serviço disponível para ele. Nessa etapa foram utilizadas diferentes fontes da literatura, tais como Targino (2000), Kuhn (2006), Cunha e Cavalcanti (2008), Alvarado e Oliveira (2008), além das fontes de metodologia científica, Severino (2007), Marconi e Lakatos (2010) e Silva (2010), e de fontes oficiais que estabelecem diretrizes para os programas de mestrado e doutorado. Em resumo, as necessidades de informação do pesquisador no exercício da profissão foram identificadas como: a) temas que já foram investigados, para apreender o que já foi feito e para avançar em seu conhecimento; b) métodos e teorias utilizados nas pesquisas, com a para ampliar o conhecimento e aprimorar teorias e métodos científicos existentes; c) resultados alcançados nas pesquisas, para entender, expandir e aplicar o conhecimento científico já conhecido, sem incorrer em retrabalhos.

Após a identificação do domínio, do usuário e suas necessidades de informação, seguiu-se com o desenvolvimento da taxonomia facetada navegacional (TAFNAVEGA), com a determinação das categorias fundamentais temáticas (CAFTE) e, depois, passando à apresentação do algoritmo utilizado para a extração de conceitos dos documentos da amostra.

3.3.2 Determinação das categorias fundamentais temáticas (CAFTE)

O conjunto CAFTE foi determinado utilizando o método da Análise de Conteúdo, com a técnica da análise categorial temática (ACT), aplicada com a proposta metodológica de Moraes (1999). Seguindo essa proposta, estabeleceu-se primeiramente o objetivo pretendido com o serviço de informação, que foi “verificar qual conjunto de conceitos é necessário para representar a estrutura textual das teses e dissertações, de forma que possa ser utilizado como categorias fundamentais temáticas na TAFNAVEGA” (MACULAN, 2011, p.102). Em seguida, para atender ao objetivo, foram executadas cinco etapas: 1) preparação das informações; 2) unitarização ou transformação do



conteúdo em unidades de conteúdo e de registro; 3) categorização ou classificação das unidades em categorias; 4) descrição; 5) interpretação.

Na primeira etapa, de preparação das informações, foram analisados todos os documentos da amostra (41 teses e dissertações), garantindo que atendesse ao critério de representatividade dentro do domínio.

O segundo passo, a etapa de unitarização ou transformação do conteúdo em unidades de conteúdo e de registro, foi um processo contínuo durante a aplicação da metodologia e do manuseio e análise dos documentos. As unidades de conteúdo (UC) e as unidades de registro (UR) foram obtidas a partir da transformação do conteúdo respeitando-se a estrutura textual dos documentos do tipo teses e dissertações (ABNT-14724, 2011) e nas seguintes fontes especializadas sobre metodologias de pesquisa: Severino (2007), Marconi e Lakatos (2010) e Silva (2010), conforme pode ser observado no QUADRO 1.

QUADRO 1 – Consolidação da estrutura textual de documentos acadêmicos

PARTE TEXTUAL (unidade de contexto)	SUBPARTES (unidade de registro)
Introdução	Tema Tempo e espaço Problema Objetivos Justificativa Metodologia Referencial teórico
Desenvolvimento	Revisão literatura Hipóteses e variáveis Fundamentação teórica Métodos e técnicas Análise e interpretação
Conclusão	Conclusão (em relação às hipóteses e objetivos)

Fonte: Baseado na ABNT-14724:2011, Severino (2007), Marconi e Lakatos (2010) e Silva (2010).

Essa consolidação foi utilizada como base que deu origem ao primeiro conjunto CAFTE, procedimento efetivado na terceira etapa, de categorização ou classificação das unidades de conteúdo (UC) e de registro (UR) em categorias, que está discriminado a seguir. Esse primeiro conjunto foi determinado *a priori*, de forma dedutiva, denominado pela técnica de análise categorial temática



(ACT) de “categorias iniciais”, resultando na seguinte codificação: C1.Tema; C2.Dimensão; C3.Ação; C4.Causa e efeito; C5.Fundamento teórico; C6.Métodos; C7.Coleta de dados; C8.Ponto de vista do autor/perspectiva; C9.Principais resultados alcançados.

Ainda nessa terceira etapa, houve a validação do conjunto de “categorias iniciais”, a partir dos critérios prescritos pela ACT de atender às características: válidas, exaustivas e homogêneas; mutuamente exclusivas; e consistentes. Essa validação foi formulada durante toda a aplicação da técnica, até o término da etapa quatro, de descrição.

Como primeiro passo da quarta etapa, houve nova validação do conjunto de “categorias iniciais”, através da descrição de cada uma das categorias, numa abordagem qualitativa, chegando-se ao seguinte conjunto e codificação, denominada pela ACT de “categorias intermediárias”: C1.Tema; C2.Ambientação; C3.Escopo; C4.Causa e efeito; C5.Fundamento teórico; C6.Métodos; C7.Coleta de dados; C8.Ponto de vista do autor/perspectiva; C9.Resultados.

Em seguida, esse conjunto de “categorias intermediárias” foi novamente validado, com uma adequação semântica e terminológica, além de reformulação da ordem lógica das categorias, o que resultou no conjunto codificado de “categorias finais”, empregado na TAFNAVEGA: C1.Tema; C2.Objeto empírico; C3.Escopo; C4.Ambientação; C5.Tipo de pesquisa; C6.Coleta de dados; C7.Métodos; C8.Fundamento teórico; C9.Fundamento histórico/contextual C10.Resultados.

Na quinta e última etapa, de interpretação, foi realizada uma análise interpretativa do resultado alcançado, também com uma abordagem qualitativa, comparando o conjunto final CAFTE com a literatura especializada da área. Essa comparação foi impetrada a partir da Norma NBR 14724:2011 e das fontes especializadas já discriminadas anteriormente, no que concerne à recomendação para a estrutura textual de teses e dissertações. A conclusão da interpretação indicou que o conjunto final CAFTE é a “representação simplificada” do conteúdo dos documentos, tal como apresentado por Marconi e Lakatos (2010, p.93). Percebeu-se que esse conjunto reflete as partes mais estáveis da estrutura textual do documento, transformando-se em um modelo do mesmo, fazendo prevalecer a ideia do todo do documento.

Concluída a determinação das dez categorias fundamentais temáticas (CAFTE) da TAFNAVEGA, passou-se para a criação da matriz para a análise conceitual, que foi obtida a partir do desenvolvimento de um algoritmo para o “processo de extração de conceitos” nos documentos. Esse algoritmo se baseou na metodologia criada por Fujita e Rubi (2006) para o Modelo de Leitura Documentária (MLD), usado pelas autoras na indexação de artigos da área das ciências biológicas.

3.3.3 A criação do algoritmo da matriz categorial temática

No procedimento de criação do algoritmo da matriz, foi desenvolvida uma sequência de tarefas, com instruções claras e bem definidas, com o objetivo de extrair os conceitos para alimentar o conjunto final CAFTE. Esse algoritmo é composto de uma tabela adaptada do Modelo de Leitura



Documentária (MLD) elaborado por Fujita e Rubi (2006) e utilizando a mesma metodologia, que consiste na formulação de três colunas, apresentadas a seguir.

A primeira coluna foi composta pelos conceitos representativos da estrutura textual dos documentos do tipo teses e dissertações, que é o conjunto final CAFTE (indicadas anteriormente como “categorias finais”). A segunda coluna acomodou os “questionamentos”, conforme sugerido pela norma NBR 12676:1992 e pelos princípios de análise do sistema de indexação PRECIS (FUJITA; RUBI, 2006). Os questionamentos representam a abordagem sistemática que foi aplicada durante a análise conceitual nos documentos. Na última coluna foram indicadas as partes da estrutura textual das teses e dissertações nas quais podem ser encontradas as respostas aos questionamentos da segunda coluna.

QUADRO 2 – Matriz categorial para trabalhos acadêmicos (teses e dissertações)

CONCEITOS CAFTE	QUESTIONAMENTOS (NORMA 12.676) e PRECIS	PARTE DA ESTRUTURA TEXTUAL
C1. TEMA	Qual o assunto de que trata o documento?	RESUMO / INTRODUÇÃO (problema, justificativa, objetivos)
C2. OBJETO EMPÍRICO	Qual o objeto empírico do estudo em questão? Qual objeto foi utilizado e/ou analisado na pesquisa?	RESUMO / INTRODUÇÃO (problema, justificativa, objetivos) / METODOLOGIA
C3. ESCOPO	O que pretende a pesquisa, de forma geral e específica, que seja relevante determinar? A que objetos a pesquisa tem intenção de atender (aprimorar, avaliar, analisar, identificar, contribuir, etc.)?	RESUMO / INTRODUÇÃO (problema, justificativa, objetivos)
C4. AMBIENTAÇÃO	O tema, objeto empírico e/ou ação são considerados no contexto de um lugar específico ou ambiente?	RESUMO / INTRODUÇÃO (problema, justificativa, objetivos) / METODOLOGIA



CONCEITOS CAFTE	QUESTIONAMENTOS (NORMA 12.676) e PRECIS	PARTE DA ESTRUTURA TEXTUAL
C5. TIPO DE PESQUISA	Tendo em vista a natureza (básica, aplicada), a abordagem (quantitativa, qualitativa), os objetivos (exploratória, descriptiva, explicativa) ou os procedimentos (bibliográfica, experimental, documental, estudo de caso, pesquisa-ação, levantamento, <i>expost-facto</i> , participante), quais as classificações podem tipificar a pesquisa realizada?	RESUMO / METODOLOGIA / RESULTADOS / DISCUSSÃO DE RESULTADOS
C6. COLETA DE DADOS	Quais instrumentos específicos (questionários, entrevistas, registros áudios-visuais, coleta de documentos, etc.) foram utilizados para realizar a ação?	RESUMO / METODOLOGIA / RESULTADOS / DISCUSSÃO DE RESULTADOS
C7. MÉTODOS	Quais modos específicos foram utilizados para realizar a ação (por exemplo, técnicas ou métodos para tratamento dos dados, que podem ser do tipo: modelagem estatística, análise estrutural, codificação, análise de conteúdo, indexação, análise semiótica, retórica ou de discurso)?	RESUMO / METODOLOGIA / RESULTADOS / DISCUSSÃO DE RESULTADOS
C8. FUNDAMENTO TEÓRICO	Houve alguma corrente ou abordagem teórica específica (teorias, hipóteses, pressupostos, etc.) utilizada em função da natureza do objeto a ser pesquisado e dos objetivos pretendidos, que foram descritos na pesquisa?	RESUMO / REVISÃO DE LITERATURA / FUNDAMENTAÇÃO
C9. FUNDAMENTO HISTÓRICO/ CONTEXTUAL	Quais temas foram tratados e revisados, a partir de pesquisa bibliográfica, para contextualizar o tema pesquisado de forma profunda e consistente?	RESUMO / REVISÃO DE LITERATURA / FUNDAMENTAÇÃO
C10. RESULTADOS	Quais pontos a pesquisa alcançou, levando em consideração os objetivos propostos? Houve formulação ou reformulação de teoria, criação de um método ou de um produto?	RESUMO / RESULTADOS / DISCUSSÃO DE RESULTADOS / CONCLUSÕES

Fonte: Adaptado do MLD de Fujita e Rubi (2006).



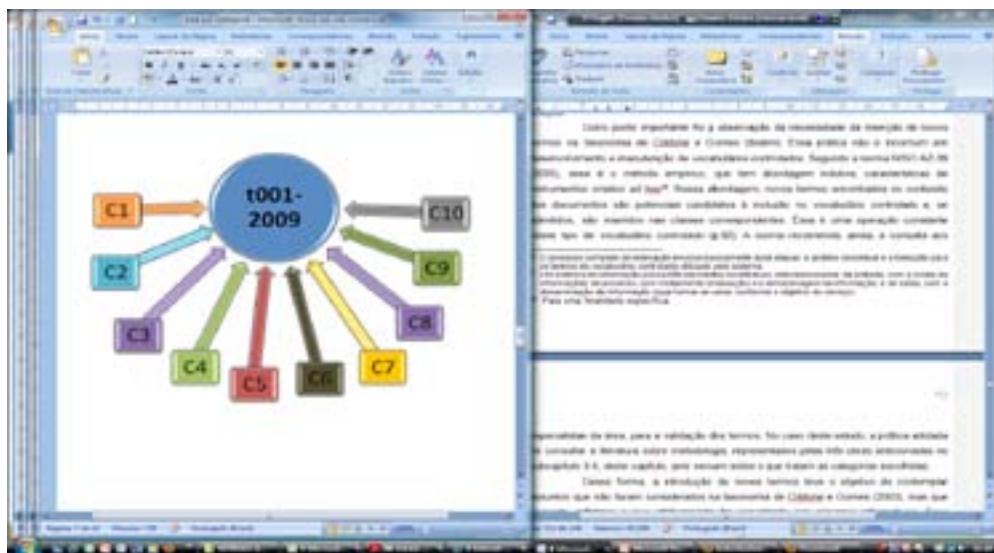
Com essa matriz (QUADRO 2) foi possível sistematizar a análise conceitual nas teses e dissertações, orientando e facilitando a extração de conceitos que representassem o conteúdo dos documentos, cujos termos ainda têm o papel de ser um recurso navegável.

3.4 Apreciação dos resultados alcançados

Para verificar a qualidade do algoritmo desenvolvido para a extração de conceitos foi realizado um teste que constituiu-se na análise conceitual dos resumos informativos dos documentos do *corpus* (41 documentos) (FIG. 1) e na implementação tecnológica (FIG. 2), com o objetivo de demonstrar a visualização da navegação pela “taxonomia facetada navegacional” (TAFNAVEGA), em meio digital.

Na análise conceitual nos resumos informativos, cada resumo foi indexado pelo conjunto final CAFTE (FIG. 1).

FIGURA 1 – Processo de extração de conceitos dos resumos informativos do *corpus*



Fonte: Elaborado pela autora.

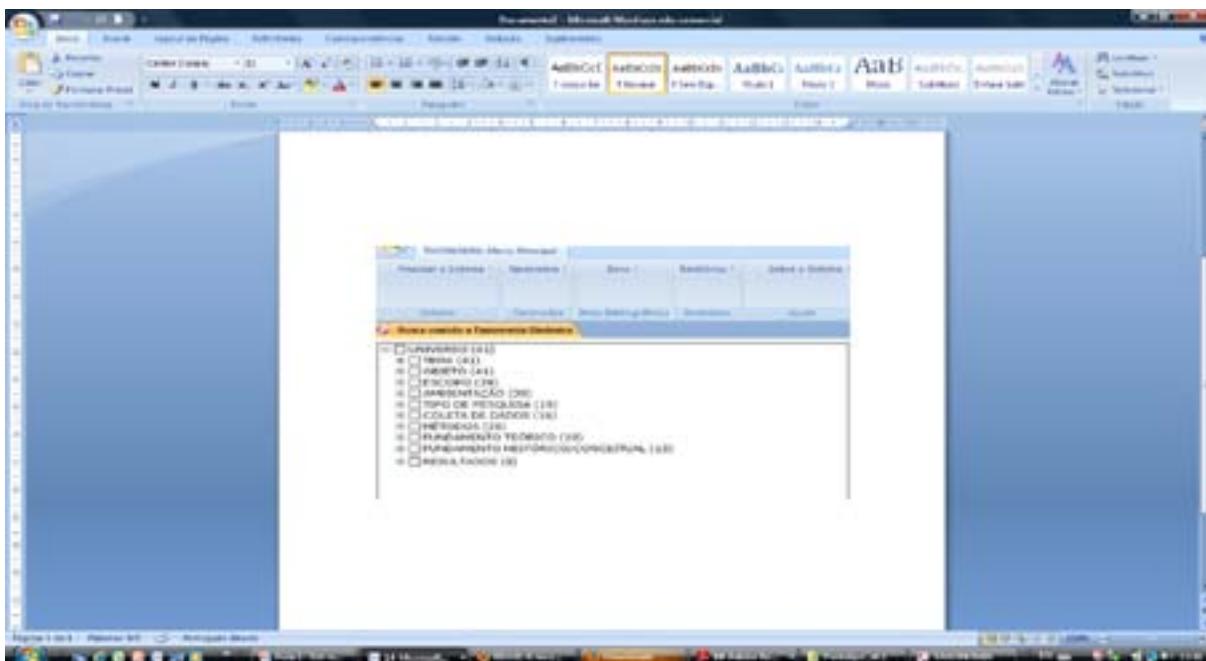
Com essa atividade foi possível estruturar o texto semi-estruturado das teses e dissertações. Esses textos são considerados “semi-estruturados” porque apresentam elementos estruturados, como autor, data, título da obra, títulos das seções, além de elementos não estruturados, como é o caso do resumo informativo e do conteúdo textual. Atribuir diferentes dimensões ao documento, a partir das categorias fundamentais, facilitou representar seu conteúdo de forma mais detalhada, pois permitiu segmentar as partes estruturais do texto, interligando seus teores.

Ao final da análise, foram obtidos 168 diferentes termos de indexação, que alimentaram as CAFTE, categorias utilizadas na etapa da implementação tecnológica (FIG. 2). A fase tecnológica foi elaborada em conjunto com Pontes (2010), dentro do laboratório experimental do “Protótipo Mapa Hipertextual” – MHTX (LIMA, 2004). Pontes (2010) desenvolveu, em seu projeto de qualificação, um aplicativo tecnológico para implementar e avaliar mecanismos para a representação e organização



do acervo de uma biblioteca digital de teses e dissertações, com base na taxonomia facetada criada na pesquisa ora apresentada neste artigo.

FIGURA 2 – Mecanismo de busca utilizando a taxonomia facetada navegacional



Fonte: Extraído do protótipo desenvolvido por Pontes (2010).

Nota-se que o UNIVERSO está representado pelos 41 documentos tratados na pesquisa de Maculan (2011), juntamente com as dez categorias fundamentais temáticas (CAFTE) utilizadas para classificá-los, que aparecem no primeiro nível da estrutura facetada. Ao lado direito de cada CAFTE está indicado o número de documentos associados a elas.

Esclarece-se que foi utilizada a Taxonomia da Ciência da Informação, desenvolvida por Oddone e Gomes (2003) como instrumento de tradução para os termos indexadores. O uso dessa taxonomia se restringiu à função de controle do vocabulário, com o objetivo de permitir a compatibilização da linguagem natural empregada nos diferentes textos, principalmente solucionando problemas de sinônimas e ambiguidades.

Um importante resultado alcançado foi observar quais tipos de informações e inferências são passíveis de serem obtidas a partir da indexação das teses e dissertações utilizando a matriz. Percebeu-se que é possível coletar, por exemplo, dados sobre: 1) os temas mais recorrentes nas pesquisas dentro da área; 2) quais objetos empíricos vêm sendo empregados, demonstrando as particularidades e interesses dos estudos, buscando entender suas motivações; 3) quais metodologias de investigação e quais abordagens, combinações de métodos e técnicas estão sendo priorizados na área; 4) os tipos de pesquisas que prevalecem; 5) os instrumentos de coleta de dados mais utilizados, o que pode indicar o grau de envolvimento dos pesquisadores nas pesquisas.

Ademais, permitiu identificar problemas tais como de resumos incompletos (95% do total), falta



de palavras-chave (25% do total) ou palavras-chave insuficientes e/ou inadequadas na representação do conteúdo das pesquisas. Além desses, foram verificados outros problemas nos resumos informativos, tais como de sinonímia na atribuição de palavras-chave, falta de objetividade e itens pré-textuais inseridos após o sumário.

Essa constatação induziu a outra verificação que foi a do uso da matriz no auxílio à elaboração de resumos informativos, uma vez que pode orientar a sistematização das informações que devem ser apresentadas em um resumo completo e de qualidade.

4 CONSIDERAÇÕES FINAIS

Todo pesquisador é também um membro da comunidade científica, que busca a aprovação de seus pares e a visibilidade acadêmica. Em sua atividade profissional, para sua produção de pesquisa, segue procedimentos metodológicos ordenados e sistemáticos. Uma pesquisa deve, numa situação processual ideal, ser construída a partir de um conhecimento anterior. Para isso, a disponibilidade e a facilidade de acesso aos detalhes de estudos já realizados anteriormente é crucial, tornando a produção científica mais eficiente e menos redundante. Acredita-se, assim, que o avanço da ciência pode ser maximizado pela comunicação do conhecimento já edificado.

Toda pesquisa científica é baseada em aportes teóricos sobre o tema investigado, aliado ao conhecimento sobre o “estado da arte” daquilo que deseja averiguar, tendo informação sobre os métodos e técnicas que são aplicáveis ao seu problema de pesquisa.

A característica semi-estruturada do conteúdo textual dos documentos do tipo teses e dissertações implica em uma dificuldade para a recuperação das informações neles contidas. As necessidades informacionais de um pesquisador, na vigência de sua investigação, incluem distintos elementos que perpassam dados sobre resultados de pesquisa, tais como teorias, métodos, instrumentos de coleta de dados ou tipos de pesquisa.

Conscientes disso, ao implantar bibliotecas digitais especializadas, de teses e dissertações, as instituições de ensino devem agregar valor ao serviço oferecido disponibilizando, além de acesso digital aos documentos, recursos que possam facilitar buscar por informações específicas e detalhadas dos mesmos. Atualmente, as BDTD possuem grande quantidade de informações que podem dar subsídio ao pesquisador, mas nem sempre elas são fáceis de explorar, localizar e recuperar.

A TAFNAVEGA foi criada a partir da ordenação lógica da estrutura textual de trabalhos acadêmicos, permitindo mapear e organizar o conteúdo considerado relevante para o pesquisador, em sua atividade profissional. Com a implementação desse recurso como mecanismo para a estruturação das informações contidas nas teses e dissertações é possível disponibilizar os documentos digitalmente e também permitir a exploração de seus conteúdos. Isso pode ser conseguido através da navegação pelas categorias e facetas da ferramenta, facilitando a busca e, consequentemente, aprimorando a recuperação da informação. A melhora é alcançada porque o usuário pode navegar pelos elementos da taxonomia facetada, escolher um aspecto sob o qual pretende efetuar sua busca e localizar a



informação desejada, sem que, em tempo algum, encontre um resultado vazio.

O mecanismo possui um conjunto de categorias fundamentais temáticas, que agrupam e interrelacionam os diferentes conteúdos dos documentos. Cada categoria é composta por uma hierarquia (taxonomia) de facetas e subfacetas, que descreve os documentos sob uma dimensão específica.

O algoritmo desenvolvido para a extração dos conceitos que alimentam as categorias fundamentais temáticas (CAFTE) da TAFNAVEGA, é utilizado no processo de indexação do recurso informacional, tornando-o sistematizado, minimizando a subjetividade que sempre envolve a atividade de análise e indexação.

Por conseguinte, a busca realizada a partir de uma taxonomia facetada navegacional permite aos usuários explorar uma grande quantidade de informações, de forma flexível, através da navegação, de modo que a sobrecarga informacional possa ser poupada.

Sendo assim, o papel fundamental da TAFNAVEGA é guiar o usuário na busca, explicitando as informações referentes ao conteúdo que realmente está disponível no banco de dados.

Abstract: This paper presents the results of a research that investigated the use of navigational faceted taxonomy (TAFNAVEGA) as an instrument of organization of information in order to give consistency and standardization to informational contents of university papers. Its aim was to facilitate the access to information by the navigation through the terms of the faceted structure. The research environment was one digital library of thesis and dissertations of a specific information unit and it had as input the thesis and dissertations of the line of research *Organization and Use of Information*. The TAFNAVEGA is composed for ten thematic fundamental categories (CAFTE) and for an algorithm for concepts' extraction called thematic categorical for academic works. The methodology included the use of theory of domain analysis to identify the domain, the BDTD user, and his information needs, as well as the Model of Documentary Reading, by Fujita and Rubi (2006) for the development of the category matrix for university works. The method of content analysis, with the technique of the thematic category analysis, was used for the fundamental thematic categories creation. The results demonstrated that the mechanism TAFNAVEGA can facilitated the task of exploration, search and retrieval of documents' content, giving access to data such as theories, methods and instruments of data gathering. The conclusion indicates that this instrument helps the researcher in its professional activity and allows to greater visibility to the available content in the digital library of thesis and dissertations without information overloading.

Keywords: Navigational faceted taxonomy. Faceted navigation. Search and information retrieval. Digital library of thesis and dissertations. Category matrix for university papers.

REFERÊNCIAS

ALVARADO, Rubén U.; OLIVEIRA, Marlene. A Comunidade científica da Biblioteconomia e Ciência da Informação brasileira. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, v. 3, n. 2, 2008. Disponível em: <<http://revista.ibict.br/pbcib/index.php/pbcib/article/view/1051>>. Acesso em: 12 jan. 2011.



ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676:1992; NBR 1052:2002; NBR 6023:2002; NBR 6024:2003; NBR 6027:2003; NBR 6028:2003; NBR 12225:2004; NBR 14724:2011; Métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro: ABNT.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. *Modern Information retrieval*. New York: Addison-Wesley, 1999.

BONILLA, Maeia Helena Silveira. *Escola aprendente: para além da sociedade da informação*. Rio de Janeiro: Quartet, 2005.

BORKO, H.; CHATMAN, S. Criteria for acceptable abstracts: a survey of abstractor's instructions. *American Documentation*, v.14, p.149-160, 1963.

BORKO, Harold. Information science: what is it? *American Documentation*, Washington, v. 19, n. 1, p. 3-5, Jan. 1968.

BROUGHTON, Vanda. The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings: New Information Perspectives*, v. 58, n. 1-2, p. 49-72, 2006.

BUCKLAND, Michel. Information as thing. *Journal of American Society of Information Science*, v. 42, n. 5, p. 351-360, 1991.

CONKLIN, Jeff. Hipertext: an introduction and survey. In: GREIF, Irene (Ed.). *Computer-supported cooperative work: a book of reading*. San Mateo, CA: Morgan Kaufmann, 1988. p. 423-476.

CONWAY, Susan; SLIGAR, Char. Building taxonomies. In: _____. *Unlocking knowledge assets*. Redmont: Microsoft Press, 2002. cap. 6.

CUNHA, M. B. da; CAVALCANTI, Cordélia R. de O. *Dicionário de Biblioteconomia e Arquivologia*. Brasília, DF: Briquet de Lemos, 2008.

ELIEL, Regiane A. Institucionalização da ciência da informação no Brasil: estudo da convergência entre a produção científica e os marcos regulatórios da área. *TransInformação*, Campinas, v. 20, n. 3, p. 207-224, set./dez. 2008. Disponível em: <http://biblioteca.ricesu.com.br/art_link.php?art_cod=6347>. Acesso em: 12 abr. 2010.

FIGUEIREDO, N. M. Da necessidade de promover o uso da informação. *Ci. Inf.*, v.6, n.1, p.75-79, 1987.

FUJITA Mariângela S. L.; RUBI, Milena P. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *DataGramZero – Rev Ci. Inf.*, Rio de Janeiro, v. 7, n. 3, jun. 2006. Disponível em: <http://www.dgz.org.br/jun06/Art_04.htm>. Acesso em: 22 abr. 2010.

FURTADO, José Afonso. Hipertexto revisited. *Letras de Hoje*, Porto Alegre, v. 45, n. 2, p. 31-55, abr./jun. 2010.

GIL LEIVA, Isidoro; FUJITA, Mariângela S. L.; RUBI, Milena P. Consistência na indexação em bibliotecas universitárias. *TransInformação*, Campinas, v. 20, n. 3, p. 233-253, set./dez. 2008.



HJORLAND, B. Domain analysis in information science: eleven approaches: traditional as well as innovative. *Journal of Documentation*, v. 58, n. 4, p. 422-462, 2002.

KUHN, Thomas S. *A tensão essencial*. 9. ed. Lisboa: Edições 70, 2006.

LIMA, G. A. B. O. *Mapa hipertextual (MHTX) um modelo para a organização hipertextual de documentos*. 2004. 199f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

MACULAN, Benildes C. M. dos S. Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos. 2011. 185f. Dissertação (Mestrado em Ciência da Informação), Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, 2011.

MALHOTRA, N. K. *Pesquisa de Marketing: uma orientação aplicada*. 3. ed. Porto Alegre: Bookman, 2001.

MARCONI, M. A.; LAKATOS, Eva Maria. *Fundamentos de metodologia científica*. 7. ed. São Paulo: Atlas, 2010.

MORAES, Roque. Análise de Conteúdo. *Revista Educação*, Porto Alegre, v. 22, n. 37, p. 7-32, 1999.

NELSON, Theodor Holm. Xanalogical structure, needed now more than ever: parallel documents, deep links to content, deep versioning, and deep re-use. *ACM Computing Surveys*, v. 31, n. 4, Dec. 1999.

NOGUEIRA, Maria do Carmo de C. *Artigos de periódicos: produção de pesquisadores de Ciência Espacial do INPE*. 1997. 149f. Dissertação (Mestrado em Biblioteconomia – Planejamento e Administração de Sistemas) – Pontifícia Universidade Católica de Campinas, Campinas, 1997.

ODDONE, Nanci; GOMES, Maria Yêda F. S. de F. Uma nova taxonomia para a ciência da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., Belo Horizonte, 2003. *Anais...* Belo Horizonte: ECI/UFMG, 2003.

PONTES, Flávio. Organização do conhecimento em bibliotecas digitais de teses e dissertações: uma abordagem baseada na classificação facetada. 2010. Qualificação de Tese (Doutorado em Ciência da Informação), Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, 2010. (Pesquisa em fase de finalização).

RABELLO, Rodrigo. *A face oculta do documento: tradição e inovação no limiar da Ciência da Informação*. 2009. 331f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Federal Paulista, São Paulo, 2009.

SACCO, Giovanni Maria; TZITZIKAS, Yannis (Ed.). *Dynamic taxonomies and faceted search: theory, practice and experience*. Berlin: Springer, 2009.

SEVERINO, Antônio Joaquim. *Metodologia do trabalho científico*. 23. ed. rev. atual. São Paulo: Cortez, 2007.



SILVA, Antônio Carlos R. da. *Metodologia da pesquisa aplicada à contabilidade: orientações de estudos, projetos, artigos, relatórios, monografias, dissertações, teses*. 3. ed. São Paulo: Atlas, 2010.

SOUZA, Vanessa I. de. *Indexação: teorias e práticas do corpo indexador da Universidade Federal do Rio Grande do Sul*. 2009. 36f. Monografia (Especialização em Gestão de Bibliotecas Universitárias) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.

TARGINO, M. G. Comunicação científica: uma revisão de seus elementos básicos. *Revista Informação & Sociedade: Estudos*, João Pessoa, v. 10, n. 2, 2000.

TUNKELANG, Daniel. *Faceted search*. Carolina do Norte: Morgan e Claypool, 2009.

TZITZIKAS, Y.; ANALYTI, A. Faceted taxonomy-based information management. In: INTERNATIONAL CONFERENCE ON DATABASE AND EXPERT SYSTEMS APPLICATIONS, 18., 2007, Regensburg. *Proceedings...* Linz: DEXA Society, 2007.

VIEIRA, Simone B. Indexação automática e manual: revisão de literatura. *Ci. Inf*, Brasília, DF, v. 17, n. 1, 1988.