

# MODELO PARA ANÁLISE CONCEITUAL DE TESES E DISSERTAÇÕES COM VISTAS À CRIAÇÃO DE TAXONOMIA FACETADA

*Benildes Coura Moreira dos Santos Maculan\**  
*Gercina Ângela Borém de Oliveira Lima\*\**

## RESUMO

Este artigo apresenta o desenvolvimento de um Modelo de Leitura Técnica para sistematização da análise conceitual em documentos do tipo teses e dissertações, visando à identificação dos conceitos representativos de seu conteúdo. É produto de uma pesquisa em andamento que tem como objetivo criar uma taxonomia facetada para navegação, como interface de busca, a fim de facilitar a representação, sistematização e posterior recuperação de informações em bibliotecas digitais de teses e dissertações. A fundamentação teórico-metodológica para o desenvolvimento do modelo utilizou os estudos de Fujita e Rubi (2006), que criaram um modelo de leitura para análise de artigos da área biológica. As autoras basearam o modelo proposto na Norma I2.676, da Associação Brasileira de Normas Técnicas (1992), na metodologia sugerida por Kobashi (1994), para indexação de documentos. Ainda, seguiram o sistema PRECIS (FUJITA, 2003), para identificação de conceitos. Além disso, dois elementos foram levados em consideração: 1) o usuário final, identificado como a comunidade científica, por meio da teoria de análise de domínio; 2) a estrutura textual de documentos do tipo teses e dissertações, baseada na Norma I4724:2005. O uso do modelo auxiliou na análise conceitual e minimizou a ocorrência de problemas de inconsistência na indexação dos documentos, apontados pelo resultado parcial da pesquisa.

**Palavras-chave:** Modelo de Leitura Técnica. Análise conceitual. Indexação de documentos. Recuperação de informação.

\* Mestre em Ciências da Informação pela Universidade Federal de Minas Gerais (Brasil). Doutoranda no Programa de Pós-Graduação em Ciência da Informação, na Escola de Ciência da Informação da Universidade Federal de Minas Gerais.  
E-mail: benildes@gmail.com.

\*\* Doutorado em Ciências da Informação pela Universidade Federal de Minas Gerais. Professora Adjunta da Universidade Federal de Minas Gerais.  
E-mail: glima@eci.ufmg.br.

## I INTRODUÇÃO

A indústria do conhecimento impõe à informação a condição de insumo essencial da sociedade, capaz de impulsionar investimentos em pesquisas para potencializar a geração de lucro. Além disso, desperta a ambição do prestígio científico entre seus pares. Em virtude desses fatos, os cientistas sempre se preocupam em difundir suas descobertas, e a forma dessa difusão é

desenvolvida e adaptada de acordo com as tecnologias disponíveis em cada época. E, na contemporaneidade, isso vem acontecendo recorrentemente em meio digital, exigindo maior aprimoramento nos serviços prestados pelos intermediários da informação, principalmente em coleções especializadas.

Outro aspecto importante refere-se ao fato de que a comunicação científica é tema de significativo interesse na Ciência da Informação (CI), pois é ela que permite perceber

o fluxo informacional e as diferentes abordagens teóricas, epistemológicas e práticas que caracterizam uma área do conhecimento. Além disso, possibilita um aumento da visibilidade institucional, legitimando e promovendo maior prestígio, o qual contribui, sobretudo, para a aquisição de financiamentos para suas atividades científicas.

Com o objetivo primordial de divulgar o saber acumulado nas instituições brasileiras, bem como incentivar uma melhor qualidade na produção acadêmica, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) instituiu a Portaria 13/2006. Essa Portaria estabelece a obrigatoriedade da divulgação digital das teses e dissertações defendidas nos programas de pós-graduação *stricto sensu* no país. A partir de então, notou-se o aumento na implantação de bibliotecas digitais de teses e dissertações (BDTD), através de diferentes iniciativas que buscam compartilhar e divulgar informações acadêmico-institucionais.

Contudo, pode-se considerar que na implantação desse tipo de BDTDs apenas estão sendo transferidos os documentos impressos, que se encontram nas prateleiras, para o meio digital, sem qualquer agregação de valor. Para disponibilizar os documentos geralmente estão sendo utilizadas as informações fornecidas pelo autor da tese ou dissertação, o que Alvarenga (2003) denomina de representação primária. Ou seja, utilizam-se como elementos, além do ano e nome do autor, o título, o resumo e as palavras-chave. As buscas nas BDTDs podem ser feitas a partir desses elementos elencados e também é possível o acesso aos textos completos.

Entretanto, a literatura da área indica problemas com títulos e palavras-chave geralmente inadequados, além de resumos informativos incompletos (BORKO; CHATMAN, 1963; VIEIRA, 1988; NOGUEIRA, 1997; GIL LEIVA; FUJITA; RUBI, 2008; ELIEL, 2008; SOUZA, 2009). Assim, a falta de agregação de valor ao serviço de informação prestado nas BDTDs, a recuperação do conteúdo informacional detalhado dos documentos, tais como teorias e métodos, é dificultada.

Buscando oferecer uma opção de busca para facilitar a recuperação das informações contidas nas teses e dissertações, criou-se uma taxonomia facetada para navegação com o propósito de oferecer uma interface de busca para a BDTD

da Escola de Ciência de Informação (ECI), da Universidade Federal de Minas Gerais (UFMG).

Uma “taxonomia é a coleção de termos de um vocabulário controlado, organizados em uma estrutura hierárquica [na qual] cada termo [...] está em uma ou mais relações tipo pai/filho (geral/específico) em relação a outro termo nessa taxonomia” (ANSI/NISO Z39.19, 2005, p.9). Ainda segundo a Norma, “uma taxonomia é um vocabulário controlado constituído por termos preferidos, relacionados entre si, em hierarquia ou poli-hierarquia” (*ibidem*, p.18). Ora, se há a possibilidade de poli-hierarquia, o mais adequado seria uma taxonomia facetada, que, segundo Lambe (2007), consiste na classificação do conteúdo de um documento sob diferentes dimensões (facetas). Cada faceta é uma dimensão diferente ou um campo de registro de metadado distinto, que pertence a uma categoria específica (mutuamente exclusivas).

Essa característica da taxonomia facetada pode auxiliar na recuperação do recurso informacional da BDTD, pois o conteúdo de um documento não está restrito a uma única dimensão. A taxonomia facetada oferece diferentes opções de busca ao usuário, o que facilita a recuperação de informações, uma vez que ela é composta por um conjunto de taxonomias, e em cada uma delas o domínio (documentos) está descrito sob um aspecto (dimensão) diferente (TZITZIKAS *et al.*, 2004).

No intuito de criar tal taxonomia facetada, foi necessário desenvolver um modelo para sistematizar a análise conceitual nas teses e dissertações, visando a extração de conceitos que irão alimentar a taxonomia facetada, denominado Modelo de Leitura Técnica (MLT). Este artigo, portanto, apresenta esse modelo para a indexação dos documentos do tipo teses e dissertações, que se baseia no modelo apresentado por Fujita e Rubi (2006).

Para atingir a esse objetivo, inicia-se com a aplicação da teoria da análise de domínio para a identificação do domínio, do usuário final da BDTD e de suas necessidades de informação, uma vez que para modelar um serviço de informação é preciso estabelecer esses elementos. Em seguida, apresentam-se as características dos documentos acadêmicos do tipo teses e dissertações de forma a indicar sua estrutura textual de produção. Finalmente, indica-se e descreve a metodologia aplicada no desenvolvimento do MLT.

## **2 O USO DA TEORIA DA ANÁLISE DO DOMÍNIO**

O termo “domínio”, no contexto da teoria da análise de domínio, refere-se a uma área de conhecimento (especialidade), a um conjunto literário ou, ainda, à tarefa de um grupo de pessoas trabalhando juntas num projeto ou numa organização (HJORLAND; ALBRECHTSEN, 1995).

É importante ressaltar que Hjørland foi pioneiro na formulação explícita da teoria da análise de domínio, que propõe entender as funções da informação em relação à sua comunicação dentro de uma comunidade discursiva. Para Hjørland e Albrechtsen (1995, p.400), comunidades discursivas são distintos grupos sociais sincronizados em pensamento, linguagem e conhecimento, constituintes da sociedade moderna. Antes de Hjørland, Jesse H. Shera, em início da década de 70, também já pensava a informação como originada pelo contexto social e cultural. Nesse cenário, o ponto central da análise de domínio consiste em estudar as atividades e produtos de uma comunidade discursiva, já que “instrumentos, conceitos, significados, estruturas de informação, necessidades informacionais e critérios relevantes estão refletidos nas comunidades discursivas” (HJORLAND, 2002b, p.258).

Ademais, Hjørland (2002b) acredita que para implementar critérios de indexação e recuperação da informação não se pode confiar apenas nos “estudos de usuários”, mas deve-se criá-los a partir de paradigmas, de normas metodológicas e epistemológicas. Esse mesmo autor afirma que a epistemologia permite a fundamentação teórica geral sobre a organização do conhecimento, mas não pode substituir o conhecimento do assunto, do domínio. Para Hjørland, a Ciência da Informação (CI) deve estar aberta a diferentes visões de organização e representação da informação, que estejam ordenadas com as características específicas de cada domínio em particular.

Nesse sentido, Hjørland (2002b) sugeriu modos específicos de estudar os domínios dentro da área da organização e representação do conhecimento: produção de classificações especiais, pesquisa especializada em indexação e recuperação, estudos da história da transferência de informações, estudos terminológicos,

linguagens para objetivos especiais (LSP), estudos do discurso, entre outros. Sugeriu, ainda, quatro métodos gerais para adquirir o conhecimento sobre o domínio: 1) métodos racionalistas (Ranganathan e a análise facetada); 2) métodos empíricos (bibliometria, co-citação); 3) métodos históricos; e 4) métodos pragmáticos (concentrando-se em objetivos, metas e utilidade).

Ainda segundo Hjørland (2002a), cada domínio possui uma estrutura conceitual própria, que o difere de outro domínio. A abordagem da análise de domínio alia, então, teoria e prática, uma vez que apresenta um método funcional (teoria) que tem por objetivo compreender as funções explícitas e implícitas da informação e as subjacentes nos fluxos de comunicação (na prática). Dessa forma, a análise de domínio se constitui em uma perspectiva que leva em consideração o discurso da comunidade na qual esse discurso é formulado e sua relação com a sociedade na qual é produzida.

É relevante, nesse contexto, o entendimento de que toda prática social está submergida em práticas informacionais, uma vez que qualquer interação humana implica geração, recepção e transferência de informação. Além disso, ainda devem ser consideradas a dimensão social representada pelas condições de produção da informação, os objetos utilizados, os valores e os sentidos que acompanham essas informações. A partir daí, pode-se concluir que a informação, ao mesmo tempo em que configura um elemento que permite as relações sociais, também é produzida nessas relações. Esse modo de pensar se aproxima do pragmatismo, no qual o principal foco é que a verdade reside no sentido de utilidade e do sentido prático das coisas, ou seja, qualquer ato, objeto ou proposição deve ser útil ou gerar um efeito prático. Portanto, nas práticas informacionais, é preciso ter em vista os domínios de discurso e as redes sociais nas quais essas práticas acontecem.

Aceitando-se esse caráter relacional do ser humano, apresentada pela teoria da análise de domínio, tem-se que o indivíduo é um ser constituído por relações intersubjetivas, como ser coletivo, e não como ente individual. Partindo dessa premissa, verifica-se que as relações do ser humano com o mundo (realidade) estão sempre mediadas por informações produzidas, armazenadas, transferidas e compartilhadas no

meio social (comunidade discursiva), as quais ordenam e fornecem sentido a essa relação (entre os pares).

Diante do exposto, define-se o domínio como a “comunidade científica”, pois é ela que produz e que também comunica os documentos acadêmicos do tipo teses e dissertações. Por conseguinte, para a identificação do usuário foram respondidos os questionamentos: Quem é esse usuário (comunidade discursiva) que necessita utilizar o recurso informacional disponibilizado pelas BDTDs? A quem interessam as teorias, métodos e resultados oriundos das diferentes pesquisas, em cada área de conhecimento? Para quem o saber acumulado nas teses e dissertações pode ter importância?

Conjecturando sobre essas questões, a conclusão inicial é de que todo conhecimento produzido abarca o interesse da sociedade em geral, e pode até mesmo impactar, de forma generalizada, a vida da coletividade. Entretanto, refletindo mais detidamente sobre o assunto, pode-se apreender que, em sentido mais restrito, cada resultado apresentado pelas pesquisas é importante para um ou outro segmento da sociedade. Porém, o ponto primordial e intencional da comunicação científica é a difusão aos pares (comunidade científica), pois os resultados obtidos nas investigações precisam de validação (que é também a comunidade científica).

Dessa forma, considerando-se que o objetivo primário da comunicação científica é a difusão de sua produção para aprovação pelos pares, tem-se que o usuário das publicações do tipo teses e dissertações é a própria comunidade científica.

## 2.1 O USUÁRIO: COMUNIDADE CIENTÍFICA

Uma comunidade científica, ou comunidade acadêmica, é representada por cientistas que, contemporaneamente, podem ser assim descritos:

[...] pessoa que possui treinamento, habilidade e desejo de procurar novos conhecimentos, novos princípios e novos materiais em algum campo da ciência (DICTIONARY OF SCIENTIFIC AND TECHNICAL TERMS, 1984). E entendo também como aqueles que fazem

pesquisa para tentar entender e ampliar o conhecimento científico já conhecido. São aqueles que fazem trabalho criativo na ciência e tratam de alguma maneira de ampliar o conhecimento ou de melhorar as teorias científicas existentes (STEVENSON; BYERLY, 1995) (ALVARADO; OLIVEIRA, 2008, p.14).

Pelo revelado acima, entende-se que seja qual for o propósito da pesquisa, os cientistas em geral, “independentemente de seu campo de conhecimento, têm necessidades de diferentes espécies de informação, a cada estágio da pesquisa, no processo de gerar, disseminar e usar a informação” (FIGUEIRA NETTO, 1994, p.29).

Uma comunidade científica, segundo Thomas Kuhn (2006), possui, entre os seus integrantes, interesses científicos semelhantes, uma bibliografia similar, vocabulário próprio (ainda que divergente em alguns pontos), compartilha modelos e abordagens metodológicos comuns, partindo, geralmente, do resultado de pesquisas já realizadas. Para esse mesmo autor, a comunidade é fortalecida à medida que o conhecimento produzido pode ser transferido e utilizado em novos aspectos de investigação dentro dela, intensificando o uso de suas próprias teorias, métodos e técnicas.

Sendo a informação insumo para atividades, seja de ordem pessoal, técnica ou científica, para esta última a informação é essencial. Ademais, para o pesquisador estar apto a novas descobertas é necessário que possa utilizar os resultados de pesquisas já concluídas, prática que evita os retrabalhos e a duplicação de esforços. Além disso, conforme Targino (2000, p.54), é “a comunicação científica que favorece ao produto (produção científica) e aos produtores (pesquisadores) a necessária visibilidade e possível credibilidade no meio social em que produto e produtores se inserem”. Em outros termos, entende-se que a comunicação científica dará visibilidade ao indicar a credibilidade ou o refutamento de um conhecimento desenvolvido em determinada comunidade discursiva, promovendo a consolidação dessa área ou atividade.

O meio digital tem ampliado a comunicação científica, tendo em vista que cientistas vêm realizando buscas mais corriqueiramente, procurando, assim, restringir o tempo gasto para encontrar as informações que necessitam. Com isso, verifica-se uma tentativa de adequação dos serviços prestados pelas bibliotecas a essa

nova possibilidade, moldando seus préstimos à demanda de sua comunidade específica.

Um dos meios de comunicação científica é a divulgação do conhecimento acumulado nas pesquisas em bibliotecas digitais (BD), instituídas em universidades nacionais e estrangeiras, tendo como objetivo disponibilizar sua produção técnico-científica. Em consequência da criação de distintas BDs, o acesso aos resultados dessas pesquisas tornou-se facilitado. Entre as diferentes iniciativas institucionais, cabe destacar o aumento na implantação de bibliotecas digitais de teses e dissertações (BDTD). Esse fato pôde ser observado, principalmente, após a Portaria 13/2006, instituída pela Capes, mencionada anteriormente na introdução. A obrigatoriedade da disponibilização das teses e dissertações em meio digital tem como objetivos dar amplo acesso ao conhecimento científico, como forma de divulgação da ciência em geral (SOUZA, 2006, *on-line*). Para alcançar esses fins, a Capes exige a instalação e manutenção de um banco digital com a inclusão de todas as teses e dissertações defendidas a partir de março de 2006, com acesso fácil, público, gratuito e permissão de *download* (SOUZA, 2006, *on-line*).

Para se adequarem à Portaria da Capes, as universidades têm procurado disponibilizar sua produção de teses e dissertações. Uma solução que essas instituições encontraram é a oferecida pelo IBICT, que coordena um Portal da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Pensando na interoperabilidade entre os diferentes sistemas, a BDTD-IBICT desenvolveu o padrão de metadados brasileiro (MTD-BR), compatível com os padrões *Dublin Core* e *Metadata Standard for Electronic Theses and Dissertations* (ETD-MS), adotado pela *Networked Digital Library of Theses and Dissertation* (NDLTD), que é um empreendimento internacional que disponibiliza textos completos de teses e dissertações publicadas em instituições distribuídas em vários países, inclusive nas brasileiras. Além disso, implementou o Protocolo OAI-PMH para coleta automática de metadados (*harvesting*).

Com isso, garante o compartilhamento de informações entre os distintos sistemas e, ainda, a biblioteca disponibilizada pelo IBICT atua como uma interface entre as diversas BDTDs no país. Nesse sentido, o projeto foi criado a fim de integrar os sistemas de informação de teses e dissertações existentes nas instituições de ensino e pesquisa brasileiras, estimulando o registro e a publicação desses documentos acadêmicos em

meio eletrônico, em um sistema de recuperação de informação (SRI).

Os documentos acadêmicos têm como função primordial difundir o conhecimento científico institucional, podendo ser na forma de artigos, relatórios, teses e dissertações, entre outros. Não obstante, neste artigo serão tratados os documentos acadêmicos do último tipo teses e dissertações.

### 3 DOCUMENTO ACADÊMICO DO TIPO TESES E DISSERTAÇÕES

Um “documento”, desde a concepção de Otlet, pode ser praticamente qualquer objeto de informação.

Para Paul Otlet, documento é o livro, a revista, o jornal; é a peça de arquivo, a estampa, a fotografia, a medalha, a música; é também, o filme, o disco e toda a parte documental que precede e sucede a emissão radiofônica. Ao lado dos textos e das imagens, existem objetos documentais por si mesmos (reália). São as amostras, os espécimes, modelos, fac-símiles e, de maneira geral, tudo que tenha caráter representativo em três dimensões e, eventualmente, em movimento (CUNHA, CAVALCANTI, 2008).

Para Rabello (2009, p.12), o documento

[...] faz parte do quadro conceitual da CI. Ele apresenta a materialidade e a susceptibilidade de inscrição ou registro informacional como características principais e apresenta, igualmente, uma estrutura e uma disposição que facilita e/ou propicia a identificação e a apreensão das informações nele registradas.

Esses dois conceitos de documento traduzem o expresso na norma NBR-12676:1992, que trata das regras para a prática normalizada da análise de documentos. Essa norma determina que documento é “qualquer unidade, impressa ou não, passível de catalogação ou indexação”. Acrescenta-se, ainda, o seguinte conceito estipulado em nota:

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.

Em resumo, adota-se, neste artigo, o conceito de “documento” como qualquer unidade informacional, impressa ou não, passível de catalogação ou indexação, seja qual for a sua natureza ou suporte. É considerada, ainda, a unidade informacional (informação) como tangível, pois, segundo Buckland (1991), pode ser tratada e disseminada em SRIs, por meio de três usos: 1) informação como processo, com a transmissão e comunicação de uma informação; 2) informação enquanto conhecimento, capaz de reduzir incertezas, com o uso da informação pelo indivíduo; e 3) informação na perspectiva de objeto, que pode ser atribuído a qualquer unidade informacional, a qualquer “coisa”.

Partindo desses preceitos de “documento”, os documentos acadêmicos constituem-se no conjunto de produções que formam a literatura científica, sendo que o que nos interessa neste artigo são as teses e dissertações.

Uma tese é um documento acadêmico que representa o resultado de pesquisa com tema único e com escopo bem delimitado, constituído de contribuição real para a questão levantada, com alto grau de profundidade de estudo. Deve ser elaborada sob a supervisão de um orientador, com titulação de doutor. Nesse contexto, o Conselho Federal de Educação (Parecer 977/65) afirma que “A elaboração de uma tese constitui exigência para obtenção do grau de ‘Doutor’ [...] deverá ser elaborada com base em investigação original devendo representar trabalho de real contribuição para o tema escolhido”.

Para a NBR 14724:2005, a tese é

Documento que representa o resultado de um trabalho experimental ou exposição de um estudo científico de tema único e bem delimitado. Deve ser elaborado com base em investigação original, constituindo-se em real contribuição para o especialista em questão. É feito sob a coordenação de um orientador (doutor) e visa à obtenção do título de doutor, ou similar.

A respeito da dissertação, é fato que também configura um documento acadêmico de tema único, com um escopo bem delimitado.

Seu objetivo é estudar e reunir subsídios para analisar e interpretar informações coletadas. Esse trabalho deverá testar a capacidade e habilidade de sistematização do aluno, sob a supervisão de um orientador, com titulação de doutor. Sobre a dissertação, o Conselho Federal de Educação (Parecer 977/65) salienta que:

o preparo de uma dissertação será exigido para obtenção do grau de ‘Mestre’ [...] deverá evidenciar conhecimento da literatura existente e a capacidade de investigação do candidato, podendo ser baseada em trabalho experimental, projeto especial ou contribuição técnica [...].

De acordo com a NBR 14724:2005, a dissertação constitui-se em um

Documento que representa o resultado de um trabalho experimental ou exposição de um estudo científico retrospectivo, de tema único e bem delimitado em sua extensão, com objetivo de reunir, analisar e interpretar informações. Deve evidenciar o conhecimento de literatura existente sobre o assunto e a capacidade de sistematização do candidato. É feito sob a coordenação de um orientador (doutor), visando à obtenção do título de mestre.

Como pôde ser observado, a pesquisa acadêmica inscrita nas teses e dissertações demanda conhecimento científico e metodológico, trazendo contribuições do autor ao tema escolhido, ainda que seja um pequeno aporte de conhecimento. Além disso, deve fazer uso de regras e normas técnicas para a normalização dos textos, para possibilitar a comunicação e o compartilhamento da informação.

No Brasil, a Associação Brasileira de Normas Técnicas (ABNT) é o órgão responsável pela normalização técnica e representa o país em entidades internacionais tais como a *International Organization for Standardization* (ISO) e *International Electrotechnical Commission* (IEC). A ABNT possui particular importância no âmbito de normalização de documentação, sobretudo na apresentação de documentos acadêmicos.

As diretrizes da Norma ABNT-NBR 14724:2011 são as que prevalecem, desde 17 de abril de 2011, para padronizar a elaboração e apresentação de documentos acadêmicos, dentre

eles, as teses e dissertações. Outras normas complementam-na: 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 10520:2002 (Citações), NBR 12225:2004 (Títulos de lombada).

Segundo a regulamentação, nos documentos do tipo teses e dissertações há três partes fundamentais:

[Primeira Parte: Introdução:] “devem constar a delimitação do assunto tratado, objetivos da pesquisa e outros elementos necessários para situar o tema do trabalho” (ABNT NBR 14724, 2005, p.6).

[Segunda Parte: Desenvolvimento:] “parte principal do texto, que contém a exposição ordenada e pormenorizada do assunto. Divide-se em seções e subseções, que variam em função da abordagem do tema e do método” (ABNT NBR 14724, 2005, p.6).

[Terceira Parte: Conclusão:] “parte final do texto, na qual se apresentam conclusões correspondentes aos objetivos ou hipóteses” (ABNT NBR 14724, 2005, p.6).

Dentre as diferentes seções e subseções possíveis na segunda parte (desenvolvimento) de uma tese ou dissertação, há quatro elementos principais:

Elemento 1 – Revisão de literatura: contém uma explanação sobre os estudos realizados por outros autores acerca do tema da pesquisa. Restringe-se, geralmente, às contribuições mais contundentes, situando a evolução do assunto.

Elemento 2 – Metodologia: descreve, detalhadamente, os métodos, técnicas e processos seguidos na pesquisa, explicando as hipóteses ou pressupostos, população ou amostra, os instrumentos e a coleta de dados. Os métodos inéditos desenvolvidos pelo autor devem ser justificados e suas vantagens apontadas em relação a outros autores.

Elemento 3 – Resultados: apresenta os dados obtidos durante o desenvolvimento da pesquisa, os quais podem ser oferecidos por meio de tabelas, quadros e figuras. Os diversos resultados alcançados devem ser agrupados e ordenados, sendo que, para maior clareza,

podem ser acompanhados por gráficos com valores estatísticos para maior clareza.

Elemento 4 – Discussão: pode expor uma comparação dos resultados alcançados na pesquisa com aqueles estudos já mencionados na revisão de literatura, ressaltando os aspectos que confirmem ou modifiquem, de modo significativo, as teorias estabelecidas.

Após o exposto, vê-se que o documento textual das teses e dissertações pode ser considerado padronizado em relação à sua estruturação. Sinteticamente, é possível enumerar: 1) seção inicial, com a introdução, as justificativas e os objetivos; 2) desenvolvimento, geralmente com a revisão da literatura, a metodologia, uma discussão (argumentação teórica, conceitual ou analítica), descrição dos dados, apresentação dos resultados; e 3) última seção, na qual são apresentadas as conclusões.

Quanto ao tipo de conteúdo, as teses e dissertações têm caráter didático, apresentando metodologia e resultados oriundos da análise de dados de pesquisa científica (MARCONI; LAKATOS, 1990). Usam linguagem formal e específica para a área ou domínio, através de uma terminologia própria. Entretanto, o documento textual das teses e dissertações é produzido em linguagem natural, que permite o uso de diferentes termos para um mesmo conceito (sinonímia). Esse fato gera, em consequência, dificuldade para a comunicação e compartilhamento do conhecimento registrado nesses documentos.

Considerando-se que um domínio de conhecimento possui uma terminologia própria, comungando de mesmos interesses científicos, a busca por uma padronização de sua linguagem é um mecanismo que pode promover a melhora na comunicação científica. Portanto, com a identificação estrutural de produção desse tipo de documento, foi realizada a análise conceitual nas teses e dissertações, de forma a desenvolver o Modelo de Leitura Técnica (MLT).

### **3.1 Um olhar sobre a análise conceitual**

Para se obter uma eficiente recuperação de informação, deve-se começar pela atividade de tratamento do material informacional, com o que Alvarenga denomina de representação secundária. Na tarefa de representação, é necessário, primeiramente, fazer a leitura técnica do documento,

com o objetivo de identificar os assuntos tratados e selecionar aqueles que, mais adequadamente, representam o conteúdo desse documento.

Esse processo de representação é tratado pela autora Boeghtol (1986), e corroborado por Lancaster (2004), como a identificação da *tematicidade* (*aboutness*) e *significado* (*meaning*). Para Boeghtol, a *tematicidade* é algo intrínseco ao documento, de natureza relativamente permanente, integrando a essência do mesmo, ou seja, é o conteúdo informacional. Quanto ao *significado*, pontua-se que se trata de um elemento mutável em função do local, do momento histórico, do interesse, da compreensão do usuário, entre outros fatos. Sob esse olhar, percebe-se que o processo de representação informacional se constitui em representar o conteúdo documental (*tematicidade*), estabelecendo, de antemão, os propósitos, concepções, critérios e contexto pretendidos, isto é, indicando o *significado* que se quer atingir. A partir disso, promove-se a organização dos conceitos, estabelecendo-se uma relação de comunicação entre o usuário e o sistema de recuperação da informação - SRI.

Na contemporaneidade, não cabe pensar em um serviço universal de informação, que seja apropriado para quaisquer propósitos. Nem tampouco é ponderado supor um serviço individualizado de informação, ainda que se busque oferecer opção de personalização. Assim, numa concepção pragmática, é recomendado oferecer um serviço no qual a informação seja concebida como uma construção coletiva da comunidade discursiva que a produz e usa e que, no caso em questão, trata-se da comunidade científica.

Com essas ponderações, foi desenvolvido o Modelo de Leitura Técnica (MLT) para subsidiar a análise conceitual nas teses e dissertações.

#### **4 METODOLOGIA DE DESENVOLVIMENTO DO MODELO DE LEITURA TÉCNICA (MLT)**

Toda leitura abarca uma interação entre leitor e texto. Ao construir o sentido deste, o leitor sempre se pautará em algum objetivo específico, que norteará essa atividade, seja uma mera leitura de lazer, uma busca de respaldo para refutar ou corroborar uma ideia, ou outra finalidade qualquer. Por consequência, ainda que o conteúdo de um texto possua um sentido de natureza relativamente

permanente, dois leitores, com intenções diferentes, podem extrair díspares significados dele.

Dessa forma, é fundamental compreender que um texto envolve estratégias cognitivas e metacognitivas: as primeiras são ações inconscientes do leitor e as segundas são sempre realizadas a partir de um objetivo específico, sendo que sobre elas há controle do leitor (KLEIMAN, 1993). Nesse sentido, Kleiman (1993) indica o uso de um modelo de estratégias metacognitivas para o ensino de leitura, como forma de minimizar a intuição e subjetividade utilizada nas estratégias cognitivas.

Seguindo esse pensamento apresenta-se o desenvolvimento de um Modelo de Leitura Técnica (MLT) para nortear a análise conceitual de teses e dissertações. Seu objetivo é minimizar os problemas de inconsistência que ocorrem no trabalho do indexador no momento da leitura técnica. Essa inconsistência pode ocorrer de duas formas: 1) inconsistência intraindexador, quando um profissional indexa um mesmo documento de diferentes formas, em distintas ocasiões temporais; e 2) inconsistência interindexador, na medida em que vários profissionais indexam um mesmo documento de maneiras diferentes. Na oportunidade da leitura técnica, busca-se a extração de conceitos que representem o conteúdo dos documentos indexados. A respeito disso, há, na literatura, evidência da interferência, sobre a atividade de indexação, de elementos como a subjetividade, o conhecimento prévio, a experiência profissional e a formação do indexador.

O desenvolvimento do MLT teve como finalidade minimizar tais interferências e foi elaborado a partir do princípio de divisão e da ordem lógica estrutural (textual) da produção de documentos do tipo teses e dissertações, sendo baseado em um modelo apresentado por Fujita e Rubi (2006). Nesse modelo, as autoras desenvolveram um instrumento para a análise de artigos científicos da área biológica, com a finalidade de auxiliar o indexador na complexa atividade de análise de assunto em documentos. O modelo é constituído pela “combinação das sistemáticas de identificação de conceitos para análise conceitual (primeira coluna) [metodologia proposta por Kobashi (1994)] e abordagem sistemática da Norma 12.676 (segunda coluna) com a localização dos conceitos em parte da estrutura textual (terceira coluna) [baseado no sistema PRECIS (FUJITA, 2003)]” (FUJITA; RUBI, 2006, *on-line*).

Fujita e Rubi (2006, *on-line*) alegam que “é preciso considerar a rápida evolução das técnicas de recuperação automática que acarretam um aumento da responsabilidade do indexador na determinação do assunto do documento”, chamando a atenção para o efeito negativo que “uma prática descompromissada com a representação do contexto do documento” pode ter na recuperação da informação. Nessa circunstância, o processo completo de indexação envolve basicamente duas etapas: a análise conceitual e a tradução para os termos da linguagem de indexação utilizada pelo sistema. Na primeira etapa do processo de indexação de documentos, segundo Fujita e Rubi (2006, *on-line*), a análise conceitual segue três procedimentos: 1) leitura do documento para exploração da sua estrutura; 2) identificação dos conceitos; e 3) seleção dos conceitos.

Dessa forma, no desenvolvimento do MLT, foram utilizados esses três procedimentos, de forma que, ao final, o instrumento concedesse à análise conceitual a representação do conteúdo dos documentos, de forma sistematizada, padronizada e completa.

O insumo de análise foram as teses e dissertações defendidas entre os anos de 1998 e 2009, da linha de pesquisa Organização e Uso da Informação (OUI), disponibilizadas, em julho de 2010, no banco de dados da BDTD da ECI, UFMG, totalizando 41 documentos.

#### 4.1 A metodologia aplicada no desenvolvimento do MLT

O desenvolvimento do MLT, primeiramente houve a determinação das categorias fundamentais temáticas. Nesse procedimento foi utilizado o método da Análise de Conteúdo, com a técnica da análise categorial temática (ACT), a partir da proposta metodológica de Moraes (1999).

A aplicação da proposta de Moraes (1999) incluiu cinco etapas, a saber: 1) preparação das informações, com a análise das 41 teses e dissertações; 2) unitarização ou transformação do conteúdo em unidades de conteúdo e de registro, com a contínua codificação das unidades de conteúdo (UC) e as unidades de registro (UR), que são unidades que obedeceram a estrutura textual dos documentos do tipo teses e dissertações, com base nas seguintes fontes: ABNT-14724, 2011,

Severino (2007), Marconi e Lakatos (2010) e Silva (2010), conforme QUADRO 1; 3) categorização ou classificação das unidades em categorias, com a categorização das UCs e das URs, chegando-se a três conjuntos: “categorias iniciais”, “categorias intermediárias” e “categorias finais”; 4) descrição, com a definição dos conjuntos de categorias fundamentais temáticas obtidas e sua validação; 5) interpretação, com uma análise interpretativa do conjunto “categorias finais” a partir das fontes especificadas no item 2.

**QUADRO 1 – Consolidação da estrutura textual de documentos acadêmicos**

<b>PARTE TEXTUAL (unidade de contexto – UC)</b>	<b>SUBPARTES (unidade de registro – UR)</b>
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).

As unidades de registro (UR) descritas no QUADRO 1 foram utilizadas como “categorias iniciais” e a validação, executada em todos os passos dos procedimentos, buscou verificar se o conjunto de categorias atendia às características: válidas, exaustivas e homogêneas; mutuamente exclusivas; e consistentes.

Ao final, as “categorias finais” incluíram: 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. A etapa de interpretação mostrou que esse conjunto de categorias fundamentais temáticas (CAFTE) é a “representação simplificada” do conteúdo dos documentos,

como indicado por Marconi e Lakatos (2010, p.93), pois apresenta as partes mais estáveis da estrutura textual das teses e dissertações.

A partir dessas dez categorias, passou-se para a segunda etapa da criação do Modelo de Leitura Técnica (MLT), representado por um algoritmo para o “processo de extração de conceitos” nos documentos. O algoritmo é composto por uma tabela com três colunas, com uma sequência de tarefas para a extração de conceitos para alimentar o conjunto final CAFTE.

Na primeira coluna, foi inserido o conjunto final CAFTE, já descrito anteriormente. Em seguida, na segunda coluna foram inseridos “questionamentos”, tais como os sugeridos pela norma NBR 12676:1992, que estabelece as

regras “exigíveis para a prática normalizada do exame de documentos, da determinação de seus assuntos e da seleção de termos de indexação” (NBR 12676, 1992, p.1). Além disso, foram aplicados os princípios de análise do sistema de indexação PRECIS (FUJITA; RUBI, 2006). Esses “questionamentos” devem ser utilizados como abordagem sistemática pelo indexador no momento da análise conceitual, para a identificação dos “termos de indexação”.

Finalmente, na terceira coluna, foram indicadas partes da estrutura de produção textual dos documentos, nas quais se pode encontrar as respostas aos “questionamentos” da segunda coluna. A tabela completa pode ser observada no QUADRO 2.

**QUADRO 2 – Modelo de leitura documentária para textos científicos (teses e dissertações)**

<b>COLUNA 1</b>	<b>COLUNA 2</b>	<b>COLUNA 3</b>
<b>CONCEITO (ANÁLISE CONCEITUAL)</b>	<b>QUESTIONAMENTO (NORMA 12.676)</b>	<b>PARTE DA ESTRUTURA TEXTUAL</b>
TEMA	Qual o assunto de que trata o documento?	RESUMO / INTRODUÇÃO
OBJETO EMPÍRICO	Qual o objeto empírico do estudo em questão? Qual objeto foi utilizado e/ou analisado na pesquisa?	RESUMO / INTRODUÇÃO (PROBLEMA/ JUSTIFICATIVA/ OBJETIVOS)
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)
AMBIENTAÇÃO	O tema, objeto e/ou ação são considerados no contexto de um lugar específico ou ambiente?	RESUMO / INTRODUÇÃO (PROBLEMA/ JUSTIFICATIVA/ OBJETIVOS)
TIPO DE PESQUISA	Tendo em vista a natureza (básica, aplicada), a abordagem (quantitativa, qualitativa), os objetivos (exploratória, descritiva, 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
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
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
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
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
PRINCIPAIS 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

Fonte: Maculan (2011).

Como pode ser observado, o MLT fornece uma sistematização para a análise conceitual das teses e dissertações (MACULAN, 2011), facilitando o exercício dessa atividade. O uso desse modelo reduziu os problemas de inconsistências já observados nessa atividade em diferentes pesquisas (VIEIRA, 1988; GIL LEIVA; FUJITA; RUBI, 2008; SOUZA, 2009).

Durante a análise conceitual nas teses e dissertações, mais duas colunas foram acrescentadas ao MLT, de forma a acomodar o registro dos conceitos identificados para cada categoria, extraídos do texto, conforme FIG. 1.

**FIGURA 1 - O Modelo de Leitura Técnica em uso na análise conceitual**

TIPO DOCUMENTO: (x) TESE TÍTULO NÚMERO: 001-2009				
CONCEITO (ANÁLISE CONCEITUAL)	QUESTIONAMENTO (NORMA 12.678)	PARTE DA ESTRUTURA TEXTUAL	EXTRATOS DO DOCUMENTO	TERMOS
TEMA	Qual o assunto de que trata o documento?	RESUMO / INTRODUÇÃO	Informação em saúde	Informação em saúde
OBJETO	Qual o objeto de estudo em questão?	RESUMO / INTRODUÇÃO (PROBLEMA/ JUSTIFICATIVA/ OBJETIVOS)	Análise dos registros de informação em saúde	Registros de informações em saúde
AÇÃO	O que pretende a pesquisa, de forma geral e específica, que seja relevante determinar?	RESUMO / INTRODUÇÃO (PROBLEMA/ JUSTIFICATIVA/ OBJETIVOS)	Entender as discrepâncias ocorridas no fluxo informacional, as questões de padronização, o uso desses sistemas de informação e sua capacidade de atender as necessidades informacionais demandadas no âmbito municipal, estadual e federal.	-Uso de sistemas de informação

Fonte: Elaborado pela autora.

De acordo com a FIG. 1, na quarta coluna foram inseridas as partes correspondentes às “respostas” a cada “questionamento” relativo às distintas “categorias”, extraídas dos documentos. Em seguida, na quinta coluna, foram alocados os “termos” representativos para a tradução dos “conceitos” extraídos dos documentos. Esses termos, quando considerados relevantes dentro do domínio e adequados ao objetivo proposto, podem ser considerados relevantes para alimentar a taxonomia facetada (vocabulário controlado). Aplicada em ambiente digital, esta taxonomia facetada torna-se navegável, o que se denomina “taxonomia facetada navegacional”, que pode orientar o usuário (cliente) na busca por informações (produtos e serviços) de sua necessidade, como tem sido recorrentemente utilizado em sites corporativos (Americanas.com, Amazon.com, Shoptime, Submarino, entre outros).

## 4.2 A navegação pela taxonomia facetada de busca

A navegação pela taxonomia facetada “permite combinar um ou mais valores [...] [sendo] possível obter todas as páginas escritas em português (faceta língua), escritas entre 2001 e 2002 (faceta data), do tipo texto (faceta tipo do conteúdo)” (PRAZERES *et al.*, 2006, p.2). De tal modo, o usuário está usando uma interface de busca indireta, na qual “o espaço de informação é mapeado, resultando na construção de uma taxonomia que melhor represente o mesmo” (HEUSELER, 2010, p.44-45). “O usuário pode elaborar a questão buscada progressivamente, ao mesmo tempo em que pode observar o efeito que sua escolha por uma faceta determina quais outras facetas estarão disponíveis para ele” (TUNKELANG, 2009, p.23), facilitando o compartilhamento da informação. Esse foi o ponto principal pretendido e alcançado.

## 5 CONSIDERAÇÕES FINAIS

As pesquisas em CI buscam compreender e solucionar problemas de natureza cada vez mais complexa, em muitos casos com cruzamento de técnicas e métodos, variando objetos de estudo e metodologias aplicadas, o que vai depender do olhar lançado ao objeto. Todos esses aspectos estão inscritos nas teses e dissertações aprovadas nos programas de pós-graduação, acumulando um conhecimento importante para o desenvolvimento da ciência e tecnologia.

Assim sendo, o avanço da ciência é dependente desse valioso conhecimento, fato que motivou a instituição da Portaria 13/2006 pela Capes. Porém, apenas disponibilizar os documentos em meio eletrônico não garante o acesso a esse saber. Recuperar e ter acesso a esse recurso informacional requer que ele seja devidamente indexado e disponibilizado. Nessa tarefa, a leitura e a análise conceitual do documento têm papel fundamental.

Como toda leitura pressupõe interação entre leitor e texto, para estabelecer significados e identificar conceitos que representem o conteúdo dos documentos, o trabalho do indexador envolve sempre atividade cognitiva e metacognitiva. Ainda que haja algumas iniciativas no sentido de indicar ao indexador

melhores práticas para estruturar sua leitura técnica, não há, ainda, consenso sobre a melhor forma de realizá-la. No entanto, foi percebido que o uso do Modelo de Leitura Técnica (MLT) desenvolvido para sistematizar o processo de análise conceitual das teses e dissertações mostrou-se bastante eficaz, pois orientou a tarefa de extração de conceitos representativos. Além disso, com o MLT foi possível instituir um escopo limitado (o conjunto de categorias) para a análise conceitual, estabelecendo regras coesas com tal escopo, de modo a tornar a análise mais objetiva e consistente. Isso facilita até mesmo em momentos nos quais a análise envolve assuntos que não são de domínio do indexador, norteados a identificação e seleção de conceitos.

Ademais, analisar as teses e dissertações permite identificar temas priorizados nas pesquisas e aqueles que ainda não foram contemplados. Aplicando o MLT no conjunto de documentos existentes na BDTD será possível determinar quais temas necessitam de maior atenção, indicando lacunas para pesquisas futuras. Esse mapeamento pode facilitar perceber o “estado da arte” da área, deixando antever as teorias e metodologias empregadas nas investigações, o que é importante para a sua consolidação.

No caso específico de elaboração de resumos, cujo problema foi detectado em grande parte dos documentos analisados, percebe-se que

o MLT tem uma função importantíssima. Ao serem analisados alguns dos arquivos com os registros dos conceitos extraídos dos documentos, dispostos na quarta coluna, o conjunto de “resposta”(s) a cada “questionamento” resultou em um resumo completo. Aplicando uma funcionalidade na entrada de dados da BDTD e oferecendo ao próprio autor do documento a possibilidade de entrar com os metadados de seu texto, o problema de resumos incompletos poderá ser significativamente reduzido. Para efeito de elucidação, os metadados são dados que caracterizam e representam um dado objeto. No caso deste artigo, trata-se de uma informação inteligível por computador, possibilitando identificar, descrever, localizar e gerenciar recursos digitais.

Ao final da pesquisa, após análise da totalidade dos documentos da amostra, confirma-se que o processo de comunicação científica poderá dar maior visibilidade ao conhecimento acumulado nas pesquisas, facilitando a recuperação dos conteúdos desses documentos. Haja vista que esses conteúdos nem sempre podem ser acessados, como é o caso, por exemplo, dos procedimentos metodológicos.

Reconhece-se, por fim, que a padronização assume um papel fundamental para a propagação digital do conhecimento adquirido nas pesquisas, permitindo uma coleta eficiente de metadados e contribuindo para o compartilhamento de informações.

## **MODEL FOR CONCEPTUAL ANALYSIS' ON THESIS AND DISSERTATIONS FOR THE FACETED TAXONOMY CREATION**

### **ABSTRACT**

*This paper presents the development of a Model of Technical Reading for systematization of the analysis of subject for indexing in documents, such as thesis and dissertations, for the identification of the representative concepts of its content. It is product of an ongoing research, which goal is to create a faceted taxonomy for browsing, as a search interface, aiming to aid representation, systematization and later information retrieval in thesis and dissertation digital libraries. The theoretical-methodological foundation for the development of this model adopted the studies of Fujita and Rubi (2006), in which a reading model for analysis of Biology research papers. The authors had based the proposed model on the 12.676 Norm from the Brazilian Association of Technical Norms (1992), on the methodology proposed by Kobashi (1994) for document indexation and on the PRECIS system (FUJITA, 2003) for the identification of concepts. Besides, two elements had been taken in consideration: 1) the final user, identified as the scientific community by means of the theory of domain analysis; and 2) the documents structure, such as thesis and dissertations, based in the Norm 14724:2005. The usage of this model aims in the subject analysis and minimized the occurrence of inconsistency on the indexation of documents given by the partial result of the research.*

**Keywords:** *Model Reading Technique. Conceptual Analysis. Document Indexing. Information Retrieval.*

---

Artigo recebido em 08/02/2011 e aceito para publicação em 20/12/2011

---

## REFERÊNCIAS

- ALVARADO, Rubén U.; OLIVEIRA, Marlene. A comunidade científica da biblioteconomia e ciência da informação brasileira. **Inf. & Soc.: Est.**, João Pessoa, v.18, n.1, p.13-29, jan./abr. 2008. Disponível em: <http://revista.ibict.br/pcbcb/index.php/pcbcb/article/view/1051>. Acesso em: 12 out. 2010.
- ALVARENGA, Lídia. Representação do conhecimento na perspectiva da Ciência da Informação em tempos e espaços digitais. **Enc. Bibli.**, n. 15, v. 8, 2003. Disponível em: <http://www.periodicos.ufsc.br/index.php/eb/article/view/97>. Acesso em: 15 jul. 2011.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: Métodos para análise de documentos, determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Norma NBR 14724:2005** (Elaboração de Documentos Acadêmicos); **NBR 6024:2003** (Numeração progressiva das seções de um documento escrito); **NBR 6027:2003** (Sumário); **NBR 6028:2003** (Resumo); **NBR 10520:2002** (Citações); **NBR 12225:2004** (Títulos de lombada). Rio de Janeiro.
- BEGHTOL, Clare. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. **Journal of Documentation**, v.42, n.2, p. 84-113, 1986.
- BRASIL. Ministério da Educação. CAPES. **Portaria 13**. 2006. Disponível em: <http://www.abmes.org.br/legislacao/2006/portarias.asp>. Acesso em: 22 mar. 2010.
- BUCKLAND, Michel. Information as thing. **Journal of American Society of Information Science**, v.42, n.5, p. 351-360, 1991.
- CUNHA, Murilo B. da; CAVALCANTI, Cordélia R. de O. **Dicionário de Biblioteconomia e Arquivologia**. Brasília, DF: Briquet de Lemos, 2008.
- DICTIONARY OF SCIENTIFIC AND TECHNICAL TERMS. 6<sup>th</sup> ed. New York: McGraw-Hill, 1984.
- FIGUEIRA NETTO, Silvino C. **A comunicação científica através de rede de computadores: a experiência de pesquisadores brasileiros**. 1994. Dissertação (Mestrado em Ciência da Informação) - IBICT, UFRJ, Rio de Janeiro, 1994.
- 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. **DGZ - Rev. Ci. Inf.**, v. 7, n. 3, jun. 2006. Disponível em: [http://www.dgz.org.br/jun06/Art\\_04.htm](http://www.dgz.org.br/jun06/Art_04.htm). Acesso em: 22 abr. 2010.
- FUJITA, M. S. L. **A leitura documentária do indexador: aspectos cognitivos e linguísticos influentes na formação do leitor profissional**. 321f. 2003. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) - Faculdade de Filosofia e Ciências, UNESP, 2003.
- GIL LEIVA, Isidoro; RUBI, Milena P.; FUJITA, Mariângela S. L. Consistência na indexação em bibliotecas universitárias. **TransInformação**, Campinas, v. 20, n. 3, p. 233-253, set./dez. 2008.
- HARTER, Stephen P. **Online information retrieval: concepts, principles and techniques**. Orlando: Academic Press. Inc., 1986.
- 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.
- HJORLAND, B. Epistemology and the socio-cognitive perspective in information science. **Journal of the American Society for Information Science and Technology - JASIST**, v.53, n. 4, p. 257-270, 2002b.
- HJORLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain analysis. **Journal of the American Society for Information Science - JASIS**, v.46, n.6, p. 400-425, 1995.
- KLEIMAN, A. **Oficina de leitura, teoria e prática**. São Paulo: Pontes, 1993.
- KOBASHI, N. Y. **A elaboração de informações documentárias: em busca de uma metodologia**.

1994. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1994.
- KUHN, Thomas S. **A estrutura das revoluções científicas**. 9. ed. São Paulo: Perspectiva, 2006.
- KURAMOTO, Hélio. Uma abordagem alternativa para o tratamento e a recuperação de informação textual: os sintagmas nominais. **Ciência da Informação**, Brasília, v.25, n.2, 1995.
- LAMBE, Patrick. Organising Knowledge: taxonomies, knowledge and organizational effectiveness. Oxford, England: Chandos, 2007.
- LANCASTER, F. W. Princípios de indexação. In: **Indexação e resumos: teoria e prática**. Brasília : Briquet de Lemos Livros, p.5-19, 2004.
- LANCASTER, F. W.; WARNER, A. J. Some basics of information retrieval. In: \_\_\_\_\_. **Information retrieval today**. Arlington: Information Resources, 1993. chap. 1, p. 1-19.
- MACULAN, Benildes Coura Moreira dos Santos. **Taxonomia facetada navegacional: construção a partir de uma matriz categorial para trabalhos acadêmicos**. 2011. 194f. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, 2011.
- MARCONI, M. A.; LAKATOS, E. M. **Técnicas de pesquisa**. 2. ed. São Paulo: Atlas, 1990.
- NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19-2005: guidelines for the construction, format, and management of monolingual thesauri**. 2005.
- PARECER Conselho Federal de Educação 977/65, aprovado em 3 dez. 1965. **Rev. Bras. Educ.**, Rio de Janeiro, n. 30, dez. 2005. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-24782005000300014&lng=en&nrm=isso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-24782005000300014&lng=en&nrm=isso). Acesso em: 20 out. 2008.
- PRAZERES, C. V. S. *et al.* Uma proposta de navegação na web utilizando facetas. In: ERBD/06: II Escola Regional de Banco de Dados. **Anais....** Passo Fundo: ERBD, 2006, v. 1, p. 1-6.
- 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 da Universidade Federal Paulista, UNESP, São Paulo, 2009.
- SALTON, Gerard; MCGILL, Michael J. **Introduction to modern information retrieval**. New York: Mcgraw-Hill Book Company, 1983.
- SOUZA, Allan Rocha de. A Portaria 13 de 2006 da CAPES e os direitos autorais. **Revista da Faculdade de Direito de Campos**, Ano VII, n. 8, 15 jun. 2006.
- SOUZA, Vanessa I. de. **Indexação: teorias e práticas do corpo indexador da Universidade Federal do Rio Grande do Sul**. 2009. 36f. Monografia (Especialização). Universidade Federal do Rio Grande do Sul. Faculdade de Biblioteconomia e Comunicação. Curso de Especialização em Gestão de Bibliotecas Universitárias. Porto Alegre: UFRGS, 2009.
- STEVENSON, Leslie; BYERLY, Henry. **The many faces of science: and introduction to scientists, values, and society**. Boulder: Westview, 1995.
- STRZALKOWSKI, Tomek. Natural language processing in large-scale text retrieval tasks. In: TEXT RETRIEVAL CONFERENCE, TREC-1. **Anais...** Gaithersburg, 1993. p. 173-187.
- TARGINO, M. G. Comunicação científica: uma revisão de seus elementos básicos. **Inf.&Soc.: est.**, João Pessoa, v. 10, n. 2, p. 37-85, 2000. Disponível em: <http://www.informacoesociedade.ufpb.br>. Acesso em: 10 out. 2010.
- VIEIRA, Simone B. Indexação automática e manual: revisão de literatura. **Ciência da Informação**, v. 17, n. 1, 1988.