

LUCIANE PAULA VITAL

**RECOMENDAÇÕES PARA CONSTRUÇÃO DE TAXONOMIA EM
PORTAIS CORPORATIVOS**

**FLORIANÓPOLIS, SC.
2007**

LUCIANE PAULA VITAL

**RECOMENDAÇÕES PARA CONSTRUÇÃO DE TAXONOMIA EM
PORTAIS CORPORATIVOS**

Dissertação de mestrado apresentada à Banca Examinadora do Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências da Educação da Universidade Federal de Santa Catarina, como requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Gestão da Informação

Linha de Pesquisa: Fluxos da Informação

Orientadora: Lígia Maria Arruda Café, Dr^a.

**FLORIANÓPOLIS, SC.
2007**

V836m

VITAL, Luciane Paula

Recomendações para construção de taxonomia em portais corporativos/ Luciane Paula Vital; orientadora Ligia Maria Arruda Café. – Florianópolis, 2007.
113f.

Dissertação (Mestrado) – Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Ciência da Informação, 2007.

1. Taxonomia. 2. Portais corporativos. 3. Organização da informação. 4. Representação do conhecimento. 5. Modelos de representação. I. Café, Ligia Maria Arruda. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Informação. III. Título.

CDU: 02

Catálogo na fonte por: Luciane Paula Vital CRB14/899

LUCIANE PAULA VITAL

**RECOMENDAÇÕES PARA CONSTRUÇÃO DE TAXONOMIA EM
PORTAIS CORPORATIVOS**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências da Educação da Universidade Federal de Santa Catarina em cumprimento ao requisito parcial para a obtenção do título de Mestre em Ciência da Informação.

APROVADA PELA COMISSÃO EXAMINADORA
EM FLORIANÓPOLIS, 01 DE JUNHO DE 2007.

Profa. Dra. Lígia Maria Arruda Café (Orientadora)

Profa. Dra. Marisa Brascher Basilio Medeiros - UnB

Prof. Dra. Miriam Figueiredo Vieira da Cunha – PGCIN/UFSC

AGRADECIMENTOS

A Deus, inteligência suprema, pela possibilidade de concluir mais uma etapa;

Ao Programa de Pós-Graduação em Ciência da Informação (PGCIN) da Universidade Federal de Santa Catarina pela oportunidade de desenvolvimento da pesquisa;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento concedido;

À minha orientadora, Ligia Café, por ter acreditado em mim e se mostrado uma pessoa paciente, empenhada e muito competente;

Aos professores do PGCIN, em especial a Profa. Miriam Vieira da Cunha, a Profa. Aline França de Abreu e a Profa Marília Damiani Costa, pelos novos olhares à pesquisa;

Aos professores que aceitaram participar da banca de defesa da dissertação, principalmente à Profa. Marisa Brascher Basilio Medeiros que se deslocou de Brasília a Florianópolis;

Ao Prof. Francisco das Chagas de Souza pelas proposições e sugestões significativas durante a disciplina de Pesquisa em Ciência da Informação;

Aos professores do Departamento de Ciência da Informação da UFSC;

À toda minha família pelo apoio incondicional, em especial à minha mãe;

À minhas amigas Gabriela e Vivian pelas conversas e troca de confidências durante o mestrado;

Aos meus amigos, pela torcida e compreensão com os sucessivos 'nãos';

Aos amigos da turma de 2005, Liliane, Cátia, Sonali, Elizangela, Johnny, Hermes, Gorete, Silvana, Elda, Marília e turma 2006 pela convivência, experiências e conhecimentos compartilhados;

Aos meus alunos da graduação da UFSC por terem me feito aprender muito mais que ensinar;

Aos funcionários do CED, em especial à Cecília da secretaria do PGCIN e à Vanda da secretaria do CIN;

A todos que, direta ou indiretamente, contribuíram para o desenvolvimento dessa dissertação, meu muito obrigada!

Se eu pudesse deixar algum presente a você, deixaria aceso o sentimento de amar a vida dos seres humanos. A consciência de aprender tudo o que foi ensinado pelo tempo a fora. Lembraria os erros que foram cometidos para que não mais se repetissem. A capacidade de escolher novos rumos. Deixaria para você, se pudesse, o respeito aquilo que é indispensável. Além do pão, o trabalho. Além do trabalho, a ação. E, quando tudo mais faltasse, um segredo: o de buscar no interior de si mesmo a resposta e a força para encontrar a saída (*Mahatma Gandhi*).

RESUMO

VITAL, Luciane Paula. **Recomendações para construção de taxonomia em portais corporativos**. 2007. 113f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis, 2007.

A gestão do conhecimento nas organizações, processo que envolve a geração, a codificação, coordenação e transferência do conhecimento, é imprescindível diante da competitividade da sociedade atual. No nível organizacional, este processo utiliza diferentes ferramentas, dentre elas os portais corporativos, que surgem como uma forma de unificar informações por um único ponto de acesso, compartilhando conhecimentos e possibilitando seu desenvolvimento. Entretanto, para que um portal corporativo seja eficiente é necessária a criação de meios de acesso, busca e a publicação da informação por qualquer usuário cadastrado no portal. Neste contexto de organização e recuperação da informação em ambientes corporativos, surgem os modelos de representação do conhecimento como instrumentos de apoio a estes processos. Entre eles está a taxonomia, foco central desta pesquisa. Para que uma taxonomia venha a ser utilizada plenamente, é importante que sua construção siga métodos precisos e objetivos. No entanto, as metodologias atualmente disponíveis não apresentam orientação suficientemente detalhada, o que deixa lacunas para entendimentos indevidos na sua construção e aplicação. Diante deste quadro, analisamos nesta pesquisa as metodologias de elaboração de taxonomias disponíveis na literatura com o objetivo de propor recomendações na sua composição. Desta forma, a presente dissertação examinou os modelos de desenvolvimento de taxonomias encontrados na literatura da área de Ciência da Informação entre os anos de 2000 e 2005, objetivando identificar e propor diretrizes para sua construção, tendo como referencial a funcionalidade relacionada à categorização, identificada na literatura da área de portais corporativos. A pesquisa se configura como um estudo exploratório qualitativo. Analisamos os modelos sob a ótica da funcionalidade categorização nos aspectos de: 1) Construção da taxonomia (representação do conhecimento), nos sub-itens: estabelecimento das categorias gerais, coleta dos termos, análise dos termos selecionados, controle da diversidade de significação e construção dos relacionamentos semânticos; e 2) Aplicação da taxonomia, nos sub-itens: representação da informação (indexação ou classificação), organização da informação recuperada. Esses aspectos de análise foram definidos com base na literatura da área de Ciência da Informação e de portais corporativos. A análise resultou em recomendações para o aprimoramento de metodologias para construção de taxonomias. Sendo assim, foi possível apresentar proposições embasadas nas áreas de Classificação, Indexação, Lingüística e Documentação, cujas discussões teóricas e metodológicas auxiliam o processo de construção de uma taxonomia.

Palavras-chave: Organização da informação. Representação do conhecimento. Modelo de representação. Portal corporativo. Taxonomia.

ABSTRACT

VITAL, Luciane Paula. **Recommendations for the construction of taxonomy in the corporative portals.** 2007. 113f. Dissertation (Master in Information Science) - Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis, 2007.

The knowledge of management in the organizations - process that involves the knowledge creation, codification, coordination and transferring - is indispensable due to the current society competitiveness. In the organizational level this process uses different tools such as the corporative portals which arise as means of joining information in one access point as well as sharing knowledge and enabling its development. However, in order to a corporative portal be efficient it is necessary the creation of means of access, search, and the information publishing by any user registered in the portal. In this context of organization and information recovery on corporative environment, the models of knowledge representation arise as support instruments to these processes. Among them is the taxonomy which is the central focus of this piece of research. In order to taxonomy to be completely utilized it is important that its construction follows precise and objective methods. Yet, the currently disposable methodologies don't present sufficiently detailed orientation, which leave gaps for improper understanding about its construction and application. From this perspective, we analyzed the elaboration methodologies of taxonomies disposable in the literature proposing recommendations to its composition. Therefore, the current thesis scrutinized the taxonomy development models found in the literature of the Information Sciences area from 2000 to 2005. The objective is to identify and propose guidelines for its construction with reference to the functionality related to the categorization - which is identified in the literature of the corporative portals area. This piece of research is configured as a qualitative exploratory study. We analyzed the models under the perspective of functionality categorization in the aspects of: 1) Construction of taxonomy (knowledge representation), in the sub-items: establishment of the general categories, terms collection, selected terms analysis, control of signification diversity and the construction of semantic relationships; and 2) Taxonomy application, in the sub-items: representation of information (indexing or classification), organization of the recovered information. The definition of these aspects of analysis was based in the literature of the Information Sciences area and corporative portals. The analysis resulted in recommendations for improvement of methodologies based on the Classification, Indexing, Linguistics and Documentation areas, from which theoretical and methodological discussions support the process of construction of taxonomy.

Keywords: Information organization; Knowledge representation; Representation model; Corporative portal; Taxonomy.

RESUMEN

VITAL, Luciane Paula. **Recomendaciones para la construcción de la taxonomía en sitios Web corporativos.** 2007. 113f. Disertación (Maestría en Ciencia de la Información) - Programa de Pós-Graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis, 2007.

El conocimiento de la administración en las organizaciones - el proceso que implica la creación del conocimiento, la codificación, coordinación y transfiriendo - está imprescindible debido al espíritu competitivo actual de la sociedad. En el nivel de la organización que este proceso utiliza instrumentos diferentes tales como los sitios Web corporativos que surgen como medios de la unión información en un punto del acceso así como compartir el conocimiento y permitir su desarrollo. Sin embargo, para un sitios Web corporativos es eficiente que es necesaria la creación de medios del acceso, la búsqueda, y la información que publica por cualquier usuario registrado en el pórtico. En este contexto de la recuperación de la organización y la información en el ambiente de corporative, los modelos de la representación del conocimiento surgen como instrumentos de apoyo a estos procesos. Entre ellos es la taxonomía que es el foco central de este pedazo de investigación. Para a la taxonomía ser utilizado completamente es importante que su construcción siga los métodos precisos y objetivos. Todavía, las metodologías actualmente para tirar no presentan la orientación suficientemente detallada, que sale los espacios para la comprensión impropia acerca de su construcción y la aplicación. De esta perspectiva, nosotros analizamos las metodologías de la elaboración de taxonomía para tirar en la literatura que propone las recomendaciones a su composición. Por lo tanto, la tesis actual escrutinizó los modelos del desarrollo de la taxonomía encontraron en la literatura del área de Ciencias de la Información de 2000 a 2005. El objetivo es de identificar y proponer las pautas para su construcción con la referencia a la funcionalidad relacionada al categorización - que es identificado en la literatura del área de los sitios Web corporativos. Este pedazo de investigación es configurado como un estudio exploratorio cualitativo. Analizamos los modelos bajo la perspectiva de categorización de funcionalidad en los aspectos de: 1) la Construcción de la taxonomía (la representación del conocimiento), en los sub-artículos: establecimiento de las categorías generales, llama la colección, escogido llama el análisis, el control de la diversidad del significado y la construcción de relaciones semánticas; y 2) la aplicación de la Taxonomía, en los sub-artículos: la representación de información (indexando o la clasificación), la organización de la información recuperada. La definición de estos aspectos del análisis fue basada en la literatura del área de Ciencias de la Información y Sitios Web Corporativos. El análisis tuvo como resultado las recomendaciones para la mejora de metodologías basadas en la Clasificación, Indexación, áreas de Lingüística y Documentación, de cuál discusiones teóricas y metodológicas sostienen el proceso de la construcción de la taxonomía.

Palabras-claves: Organización de la información; Representación del Conocimiento; Modelo de la representación; Sitios Web corporativos; Taxonomía.

LISTA DE ILUSTRAÇÕES

Quadro 1: Conhecimento tácito e explícito.....	20
Quadro 2: Gerações de portais corporativos.....	25
Quadro 3: Requisitos mínimos de um portal corporativo.....	28
Quadro 4: Comparativo das funcionalidades de um portal com um SRI.....	30
Quadro 5: Tipologias de relações seqüenciais.....	63
Quadro 6: Funcionalidades dos portais relacionados à organização e representação da informação.....	70
Quadro 7: Síntese da análise das práticas de elaboração de taxonomia.....	91
Quadro 8: Revocação e precisão em relação a especificidade e exaustividade.....	102
Figura 1: A espiral do conhecimento.....	21
Figura 2: Componentes-chave da arquitetura PdCC.....	27
Figura 3: Ciclo da informação.....	33
Figura 4: Exemplo de taxonomia.....	46
Figura 5: Pesquisa com o termo ‘trial’.....	48
Figura 6: Processo geral de um software de taxonomia.....	51
Figura 7: Arquitetura – Metodologia.....	53
Figura 8: Triângulo de Ullman.....	55
Figura 9: Esquema de relações hierárquicas.....	61
Figura 10: Relações seqüenciais.....	64

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	OBJETIVOS.....	17
1.1.1	Objetivo Geral.....	17
1.1.2	Objetivos Especificos.....	17
1.2	ESTRUTURA DA DISSERTAÇÃO.....	17
2	QUADRO TEÓRICO DE REFERÊNCIA.....	19
2.1	GESTÃO DO CONHECIMENTO.....	19
2.2	PORTAIS CORPORATIVOS.....	22
2.3	MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO.....	31
2.3.1	Classificação.....	34
2.3.2	Tesouros.....	41
2.3.3	Ontologia.....	43
2.3.4	Taxonomia.....	46
2.4	RELAÇÕES SEMÂNTICAS ADOTADAS EM MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO.....	54
3	PROCEDIMENTOS METODOLÓGICOS.....	65
3.1	TIPO DE PESQUISA.....	65
3.2	ETAPAS DA PESQUISA.....	65
3.2.1	Identificação e seleção das metodologias de desenvolvimento de taxonomias.....	66
3.2.1.2	Metodologias de desenvolvimento de taxonomias.....	68
3.2.2	Identificação e seleção das funcionalidades dos portais corporativos.....	69
3.2.2.1	Funcionalidades dos portais corporativos.....	70
4	ANÁLISE DAS METODOLOGIAS DE DESENVOLVIMENTO DE TAXONOMIAS.....	74

5	RECOMENDAÇÕES PARA O DESENVOLVIMENTO DE UMA TAXONOMIA EM PORTAIS CORPORATIVOS.....	92
5.1	CONSTRUÇÃO DA TAXONOMIA.....	93
5.2	APLICAÇÃO DA TAXONOMIA.....	100
6	CONCLUSÃO.....	103
	REFERÊNCIAS.....	107

1 INTRODUÇÃO

A sociedade pós-industrial, caracterizada pelo predomínio do trabalho no setor terciário, concentra sua força na produção e distribuição de informação e conhecimento (LEMOS, 1999, p.128). Esse aspecto trouxe mudanças na estrutura social, como a passagem da produção de bens para a economia de serviços, o avanço das classes de trabalhadores técnicos, o desenvolvimento tecnológico, a automação industrial e uma nova organização produtiva baseada em informação e conhecimento. Essas transformações são algumas das que determinaram o que Lojkine (2002) chama de Revolução Informacional, que transfere para as máquinas não somente a força física do trabalhador, mas também a capacidade de tomar decisões. Por isso, é vista como causadora de um impacto maior que o da Revolução Industrial. De acordo com Lojkine (2002, p.11),

a Revolução Informacional está em seus primórdios e é primeiramente uma revolução tecnológica que se segue à Revolução Industrial. Mas é também muito mais que isto; constitui o anúncio e a potencialidade de uma nova civilização, pós-mercantil, emergente da ultrapassagem de uma divisão entre os que produzem e os que dirigem a sociedade.

Conseqüentemente, essa nova sociedade possui mão-de-obra muito mais qualificada e profissões ricas em informação. A exigência de trabalhadores intelectualmente mais qualificados aliada a um cenário competitivo determinam o conhecimento, em um movimento sistemático de criação, desenvolvimento e aplicação, como fator decisivo na sobrevivência das empresas.

A informação e o conhecimento sempre foram elementos importantes no desenvolvimento econômico de qualquer sociedade. Porém, com o desenvolvimento tecnológico e a automação dos processos produtivos, até então desenvolvidos com força de trabalho humana na sociedade industrial, passaram a constituir-se em elementos essenciais para a competitividade.

A automação atingiu, inicialmente, os processos produtivos considerados padronizados e repetitivos e,

os colaboradores passaram, então, a dedicar cada vez mais seu tempo na execução de tarefas complexas, o conhecimento e a multidisciplinaridade passam a ser fatores críticos para a execução das mesmas. As competências organizacionais e funcionais também se modificam. E a informação e o conhecimento passaram a ser os recursos capitais de produção (ABREU e ABREU, 2002).

O interesse na gestão, tanto da informação quanto do conhecimento¹, apresenta conseqüências de fatores econômicos e sociais. Passamos para um período em que a gestão cognitiva, centrada no pensar e aprender, é evidenciada. Esse período recebe várias denominações, 'Sociedade do conhecimento', 'Sociedade pós-industrial', 'Sociedade da Informação', entre outras. Essa variedade de denominações refletem, como aponta Sveiby (1998, p.31), "a atual confusão em relação à maneira com percebemos o mundo futuro". Entretanto, sociedade do conhecimento parece ser o termo que melhor define o período atual pois, segundo Fuks (2003), nessa sociedade, "uma das fontes principais de produtividade e eficiência passa a ser a inovação tecnológica (novo conhecimento)". As organizações, para manter e/ou conquistar espaço no mercado, precisam se adaptar a novos ambientes, onde o valor de produtos e serviços depende cada vez mais do percentual de inovação, de tecnologia e de inteligência a eles incorporados (CAVALCANTI; GOMES; PEREIRA, 2001).

Nesse sentido, a necessidade do gerenciamento informacional sistematizado e o desenvolvimento de ações estratégicas voltadas para a otimização do conhecimento organizacional e de seus recursos tecnológicos são considerados como pré-requisitos na sociedade que se configura.

Nos ambientes organizacionais, a troca de mensagens via correio eletrônico e a divulgação de notícias, informes e comunicados disponibilizados pelas mídias, fazem com que a existência de um filtro e de formas de organizar essas informações sejam imprescindíveis. O surgimento tanto das Intranets como de portais corporativos reflete a preocupação com esse gerenciamento. No intuito de disponibilizar e prover acesso aos recursos informacionais e ao conhecimento desenvolvido, as organizações têm adotado os portais corporativos, que possuem ferramentas eficientes para esse compartilhamento. Em ambientes informacionais menores, o gerenciamento torna-se mais eficiente, não só pelo volume de informações como por estar direcionado a necessidades específicas, auxiliando melhor o compartilhamento de informações relevantes aos usuários.

Segundo Terra e Bax (2003), os portais servem para "desbloquear a informação armazenada na empresa, disponibilizando-a aos utilizadores através de um único ponto de acesso". Eles surgem como uma forma de reunir informações por

¹ "Geração, a codificação, a coordenação e a transferência do conhecimento." (DAVENPORT, PRUSAK, 1999)

um único ponto de acesso e, mais do que isso, são ferramentas de apoio à gestão do conhecimento, cujo objetivo não é só fornecer informações, mas prioritariamente compartilhar conhecimentos e favorecer seu desenvolvimento. Quando tratamos de informações organizacionais e conhecimentos, sejam tácitos ou explícitos², torna-se imprescindível a formalização e a incorporação dos processos às rotinas diárias, o que é possível ao se adotar um portal corporativo. A partir do portal, os processos de gestão do conhecimento podem ser consolidados, disseminados e gerenciados. Em organizações do conhecimento, os dados e informações precisam estar disponíveis a todos os membros de forma integrada, rápida e eficaz. Portais corporativos possibilitam, além dessa integração, a colaboração dos próprios usuários na inclusão de novos conteúdos. Essa inclusão precisa acontecer de forma padronizada, com conceitos e termos comuns e significados compartilhados, para que os portais não se tornem um caos informacional, como a Internet se apresenta atualmente.

Um fator primordial para haver compartilhamento da informação em um portal, no processo de gestão do conhecimento, é que haja também meios que viabilizem o acesso, busca e a publicação da informação por qualquer usuário cadastrado no sistema. Nesse sentido, o tratamento das informações e a codificação dos conhecimentos³ em portais corporativos requerem o desenvolvimento e a aplicação de modelos de representação comuns a todos os usuários do sistema.

De acordo com Campos (2004, p.24), os modelos de representação da informação e do conhecimento possibilitam “[...] a elaboração de linguagens documentárias verbais e notacionais, visando à recuperação de informações e a organização dos conteúdos informacionais de documentos”.

Sendo assim, a taxonomia, um modelo de representação de informações e conhecimentos, vem sendo considerada, na literatura, como uma metodologia facilitadora tanto da organização quanto da recuperação, representação e disponibilização do conteúdo em ambientes *Web*, mais especificamente em portais corporativos. Possibilita que usuários, por meio de uma conceitualização comum, encontrem os dados procurados, disponibilizem suas informações, compartilhem não

² Nonaka e Takeuchi (1997, p.7) entendem o conhecimento tácito como “[...] altamente pessoal e difícil de formalizar”, enquanto o conhecimento explícito é definido como o que pode ser “[...] expresso em palavras e números, e facilmente comunicado e compartilhado [...]”.

³ Por codificação do conhecimento entende-se o registro do conhecimento, transformando-o em informação que possa ser disponibilizada a outras pessoas em diferentes suportes.(DAVENPORT, PRUSAK, 1999)

só informações estruturadas, mas também documentos como *e-mails* e comunicados entre outros informes que circulam diariamente em uma organização e que são propulsores de novos conhecimentos. Por meio de suas estruturas hierárquicas, os termos são alocados de maneira lógica, facilitando as buscas, a inserção e a recuperação das informações. Na economia atual, globalizada e competitiva, o conhecimento que os indivíduos desenvolvem acerca dos processos organizacionais é cada vez mais valorizado e induzido à codificação. Terra e Gordon (2002) afirmam que “taxonomias são elementos estruturantes estratégicos e centrais para negócios baseados em informação e conhecimento”. A importância e o crescente uso de taxonomias em portais corporativos provocam o interesse em aprofundar nossos conhecimentos sobre o tema; pois, apesar da crescente importância desse modelo de representação do conhecimento no contexto informacional configurado em ambientes *Web*, a literatura da área é escassa em reflexões aprofundadas sobre o processo de construção.

Em princípio, as taxonomias são construídas com base em modelos hierárquicos do conhecimento. No entanto, nem sempre seguem esse modelo no que diz respeito aos princípios lógicos da organização hierárquica. Além disso, questiona-se se apenas os modelos hierárquicos seriam suficientes para a organização da informação em portais corporativos. Talvez o problema esteja no fato de as práticas de elaboração para a construção de taxonomias não definirem com exatidão os tipos de relacionamentos semânticos possíveis e como eles devem ser estabelecidos. Buscando aprofundar a reflexão e a análise a respeito do desenvolvimento das relações semânticas, esta pesquisa tem o objetivo de analisar práticas de elaboração para o desenvolvimento de taxonomias em portais corporativos.

Considerando os argumentos citados, o problema deste trabalho é explicitado da seguinte forma:

As práticas de elaboração de taxonomias contemplam orientações precisas para a construção de um modelo de representação do conhecimento apto a ser aplicado na tradução dos conteúdos informacionais em um ambiente corporativo?

1.1 OBJETIVOS

Os objetivos da pesquisa situam-se em dois níveis: o geral e o específico.

1.1.1 OBJETIVO GERAL

Propor recomendações para o processo de construção de taxonomias em portais corporativos.

1.1.2 OBJETIVOS ESPECÍFICOS

- Caracterizar os modelos de representação do conhecimento;
- Identificar as práticas de elaboração de taxonomias disponíveis na literatura da área de Ciência da Informação no período de 2000 a 2005 em inglês, português e espanhol;
- Caracterizar as funcionalidades de um portal corporativo que estejam relacionadas à organização e representação da informação;
- Analisar as práticas de elaboração identificadas sob a luz das funcionalidades dos portais corporativos.
- Estabelecer recomendações para o processo de construção de taxonomias corporativas.

1.2 ESTRUTURA DA DISSERTAÇÃO

Com relação à estrutura formal da dissertação, o trabalho está organizado em seis partes. A primeira parte é introdutória e apresenta a justificativa, a pertinência da pesquisa e a delimitação do problema, assim como, os objetivos gerais e específicos que estimularam o trabalho.

A segunda parte, Quadro Teórico de Referência, trata do aspecto teórico da pesquisa, apresentando subsídios para um entendimento amplo e aprofundando do tema, assim como, para a interpretação e posterior discussão dos resultados.

A terceira parte, Procedimentos Metodológicos, delimita e apresenta os aspectos metodológicos da pesquisa, define o tipo de pesquisa, a técnica utilizada na coleta dos dados e na exploração do material. Assim como, apresenta os dados coletados e o detalhamento da variável de análise utilizada.

A quarta parte apresenta a análise das práticas de elaboração coletadas sob a luz da literatura, permitindo uma discussão dos resultados.

A quinta parte propõe, com base na literatura e nas práticas de elaboração analisadas, recomendações para o desenvolvimento de taxonomias em portais corporativos.

A sexta parte apresenta as conclusões realizadas e as proposições para futuros trabalhos na área, seguida pela apresentação das referências que garantem o aporte teórico, assim como, a localização do material pelos interessados.

2 QUADRO TEÓRICO DE REFERÊNCIA

A seguir é apresentado o quadro teórico que fundamenta a pesquisa e embasa a análise dos dados.

2.1 GESTÃO DO CONHECIMENTO

O sucesso da gestão do conhecimento é alcançado por meio de uma organização em constante aprendizagem, cuja premissa é a geração e o compartilhamento do conhecimento. Na literatura afim, existem vários conceitos que procuram definir a gestão do conhecimento, alguns enfatizando o aspecto humano, outros o tecnológico. Para Davenport e Prusak (1999), a gestão do conhecimento envolve a geração, a codificação, a coordenação e a transferência do conhecimento. Para Nisembaum (2002, p. 275), “trata-se do processo pelo qual apoiamos a geração, o armazenamento e o compartilhamento de informações valiosas, *insights* e experiências dentro e entre comunidades de pessoas e organizações com interesses e necessidades similares”. Baseando-se em Dudziak, Villela e Gabriel (2002), podemos concluir que “[...] a visão integradora parece ser a mais coerente”; visão esta que percebe a gestão do conhecimento com base em três dimensões: como filosofia, como processo e como objeto. No que tange à dimensão filosófica, está intrinsecamente relacionada com a cultura organizacional, com os modelos mentais definidos por Senge (1998) como “[...] imagens internas profundamente arraigadas de como o mundo ou a empresa funciona”. Como processo, segundo Dudziak, Villela e Gabriel (2002), “[...] liga-se estruturalmente às atividades e tarefas necessárias à sua coleta, organização, controle, disponibilização e uso por parte de todos os membros da organização”. Concebida como objeto, de acordo com esses autores, trata-se de algo que transforma dados em informação e, posteriormente, em conhecimento, fundamentando-se em “processos interpretativos e reflexivos”.

O fato de ela assumir várias dimensões e não estar restrita a um setor da organização dificulta a construção de uma definição precisa sobre gestão do conhecimento. Alia-se a essa complexidade, a necessidade de se entender como o conhecimento se processa nas pessoas, como é internalizado e externalizado ou como é criado. O quadro 1 mostra as principais características dos dois tipos principais de conhecimento que apóiam a gestão do conhecimento, de acordo com Nonaka e Takeuchi (1997).

<i>Conhecimento Tácito</i>	<i>Conhecimento Explícito</i>
Assistemático	Formal e sistemático
Altamente pessoal	Facilmente comunicado
Difícil de ser transferido	Facilmente compartilhado
Envolve fatores intangíveis	Tangível

Quadro 1: Conhecimento tácito e explícito

Fonte: Nonaka e Takeuchi (1997)

Nonaka e Takeuchi (1997, p.7) entendem o conhecimento tácito como “[...] altamente pessoal e difícil de formalizar”. Para os autores, esse conhecimento “[...] está profundamente enraizado nas ações e experiências de um indivíduo, bem como em suas emoções, valores ou ideais”. Sendo o conhecimento tácito altamente pessoal, sua conversão em recurso disponível para outras pessoas, o compartilhamento, é a atividade central da empresa criadora de conhecimento (NONAKA, 2000, p.32).

O conhecimento explícito é definido como o que pode ser “[...] expresso em palavras e números, e facilmente comunicado e compartilhado sob a forma de dados brutos, fórmulas científicas, procedimentos codificados ou princípios universais [...]” (NONAKA e TAKEUCHI, 1997, p.7). O processo de criação do conhecimento ocorre na interação desses dois tipos. Choo (2003, p. 203) declara que “a criação do conhecimento organizacional é, portanto, um processo que amplifica 'de maneira organizacional' o conhecimento criado pelos indivíduos e cristaliza-o como parte da rede de conhecimento da organização”.

Para podermos entender melhor essas relações, apresentamos a Figura 1 que ilustra a espiral do conhecimento de Nonaka e Takeuchi (1997). Os quatro padrões básicos de criação de conhecimento interagem sistematicamente, originando assim uma cultura organizacional do conhecimento.



Figura 1: A espiral do conhecimento
Fonte: Nonaka e Takeuchi (1997, p. 80)

Teixeira Filho (2001) apresenta alguns tipos de conversão do conhecimento baseado nos quatro padrões de criação e mostra que,

Socialização: Tácito para Tácito = Comunidade
 Externalização: Tácito para Explícito = Memória
 Combinação: Explícito para Explícito = Sistemas
 Internalização: Explícito para Tácito = Treinamento

As organizações, atuando em economias globalizadas, nas quais a informação e o conhecimento são propulsores da competitividade, deveriam trabalhar com base nesses quatro padrões de criação do conhecimento. A socialização acontece quando grupos de especialistas têm a oportunidade de interagir, seja pessoalmente, seja de maneira remota, por meio da Internet e de outras tecnologias da informação, compartilhando seus conhecimentos e fomentando novos. No processo de externalização, o conhecimento tácito é codificado; entretanto, para que isso ocorra, meios adequados devem ser disponibilizados aos indivíduos. Para a potencialização da combinação e da internalização, é necessária a disponibilização de informações, de ambientes informacionais estruturados, juntamente com uma cultura de aprendizagem, em que o conhecimento seja visto como um investimento. Para trabalhar esses quatro padrões de criação do conhecimento, os portais corporativos podem se constituir em uma ferramenta-chave que, além de disponibilizar os recursos informacionais da organização, potencializa trocas e compartilhamento de informação e conhecimentos.

Terra e Gordon (2002, p.57) explicam que “a Gestão do Conhecimento, em seu sentido mais atual, pode ser considerada o esforço para melhorar o desempenho humano e organizacional por meio da facilitação de conexões significativas”. O portal corporativo pode atuar como grande facilitador no processo de estabelecimento dessas conexões significativas, alavancando a gestão do conhecimento, seja por meio de fontes de informação especializadas, seja pela possibilidade de colaboração entre os membros da organização ou mesmo no momento de codificação e disponibilização do conhecimento.

2.2 PORTAIS CORPORATIVOS

O surgimento dos portais na Internet está relacionado ao gerenciamento da crescente massa informacional criada e disponibilizada na Web, à qual nos referimos anteriormente. No final da década de 1990, os portais podiam ser caracterizados como sistemas de busca. Dias (2001, p.52) afirma que “para reduzir ainda mais o tempo de busca na Internet e auxiliar os usuários menos experientes, vários *sites* de busca incluíram categorias, isto é, passaram a filtrar sites e documentos em grupos pré-configurados de acordo com seu conteúdo [...]”. Essa forma de organizar e disponibilizar o acesso à informação tornou-se muito popular nos chamados ‘portais públicos’, que “[...] provêm ao consumidor uma única interface à imensa rede de servidores que compõem a Internet” (DIAS, 2001). Para Pérez de Leza (2000 *apud* ARCE, PÉREZ, 2001, p.223), um portal é definido como “uma página na *Web* que agrega conteúdos e funcionalidades, organizados de tal maneira que facilitam a navegação, proporcionam ao usuário um ponto de entrada na rede com um amplo leque de opções de serviços e produtos”⁴.

O objetivo inicial dos portais, fornecimento de um ponto de acesso comum a todo o conteúdo disponibilizado na Internet, logo foi substituído por interesses primordialmente comerciais, cujo desejo é manter o usuário o maior tempo possível navegando no portal. Podemos citar como exemplos de portais públicos, o *Yahoo*

⁴ “La pagina Web que agrega contenidos y funcionalidades, organizados de tal manera que faciitan la navegación y proporcionan al usuario un punto de entrada em la Red con um amplio abanico de opciones”. (PÉREZ DE LEZA, 2000 *apud* ARCE, PÉREZ, 2001, p.223)

(<http://www.yahoo.com>), o Bol (<http://www.bol.com.br>), o Terra (<http://www.terra.com.br>) e tantos outros portais de provedores de acesso à Internet.

Os ambientes organizacionais logo se 'apropriaram' da funcionalidade do portal de ser um ponto de acesso único de informações e iniciaram, a partir das Intranets, a sua implantação. Nesse contexto, os portais podem ser concebidos como uma evolução das Intranets, pois integram vários sistemas de informação em uma plataforma única, com interface comum a todos os usuários. Com eles, a organização tem um 'ponto de encontro' quando se trata de recuperar ou disponibilizar informações, possibilitando, dessa maneira, a potencialização do compartilhamento tanto de informações quanto de conhecimentos.

Podemos afirmar que portais representam uma evolução das Intranets quando entendemos Intranet como “[...] muito mais que uma mera interligação de servidores que executam tarefas semelhantes – é um sistema corporativo que envolve computadores e pessoas” (BENETT, 1997, p.21). Assim sendo, portais apresentam as principais características das Intranets, como a integração e o compartilhamento de informações, mas de modo centralizado. Carvalho e Ferreira (2005) discutem a evolução das Intranets em direção aos portais corporativos e declaram que “a missão dos portais corporativos é acabar com as ilhas dos sistemas de informações, integrando-as em uma única aplicação que seria a porta de entrada para todos os usuários do ecossistema empresarial”. Portais utilizam metadados e linguagem XML (*Extensible Markup Language*), que possibilita a integração de dados estruturados e não-estruturados. Dias (2001); Carvalho e Ferreira (2005) afirmam que não há um consenso na literatura em torno do conceito de portais; são utilizados como sinônimos de Portais de Informações Empresariais, Portais de Negócios e Portais Corporativos.

Porém, apesar das ambigüidades, a definição de Portal de Informação Empresarial (EIP – *Enterprise Information Portal*) encontrada na literatura dá ênfase maior ao gerenciamento informacional, sem abordar de forma satisfatória o aspecto cooperativo (DIAS, 2001); e, quando trata desse aspecto, envolve o conceito de cooperação em um processo mais amplo de gestão do conhecimento. As tecnologias de EIP trabalham com os conceitos tratados anteriormente, de integração e compartilhamento de informações, tanto internas quanto externas ao ambiente organizacional. De acordo com Turban, Rainer Junior e Potter (2003, p. 233), os EIP “[...] integram aplicações de gerenciamento de conteúdo, inteligência

empresarial, depósito de dados, central de dados e gerenciamento de dados”. Centralizando o ponto de acesso às informações organizacionais, internas e externas, a empresa ganha em inteligência, otimiza o compartilhamento e a produção de novos conhecimentos, além de qualificar as tomadas de decisões.

O termo Portal de Negócios, utilizado por Eckerson (*apud* DIAS, 2001), é entendido como “um aplicativo capaz de proporcionar aos usuários um único ponto de acesso a qualquer informação necessária aos negócios, esteja ela dentro ou fora da corporação”. Essa definição de Eckerson está relacionada com a gestão da informação, cujo ponto central é a disponibilização de informações referentes aos processos organizacionais, sem necessariamente gerir o conhecimento.

Já o Portal Corporativo, no conceito de Collins (2003 *apud* CARVALHO e FERREIRA, 2005), é “[...] uma interface personalizada de recursos on-line que permite que os trabalhadores do conhecimento⁵ acessem e partilhem informações, tomem decisões e realizem ações independentemente da sua localização física, do formato da informação e do local em que ela está armazenada”. Terra e Gordon (2002) dizem que “os portais são fundamentais para organizações intensivas no uso de informação e conhecimentos, [assim como] para organizações que se comunicam e colaboram de forma intensiva com um número expressivo de agentes e atores externos (clientes, parceiros, fornecedores)”. Segundo os autores, os portais melhoram a eficiência do trabalho em equipe, o uso de informações corporativas e do ambiente externo e podem aumentar a produtividade do ‘trabalhador do conhecimento’. Ele aborda os portais corporativos dentro de um processo de Gestão do Conhecimento e os denomina ‘Portais de Conhecimento Corporativo’ (PdCC).

Eckerson (1999 *apud* DIAS, 2001) verifica uma evolução dos portais corporativos e os divide em quatro gerações (Quadro 2).

⁵ Entende-se por trabalhador do conhecimento o indivíduo “[...]que sabe selecionar, absorver informação e conhecimento onde quer que ele se encontre e com capacidade para aplicar este conhecimento em ações concretas”. (HARVARD BUSINESS REVIEW, 2000)

Geração	Categoria	Características
Primeira	Referencial	Máquina de busca, com catálogo hierárquico de conteúdo da <i>Web</i> . Cada entrada do catálogo contém uma descrição do conteúdo e um <i>link</i> . Essa geração enfatiza mais a gerência de conteúdo, disseminação em massa das informações corporativas e o suporte à decisão.
Segunda	Personalizado	O usuário, por meio de um identificador e uma senha, pode criar uma visão personalizada do conteúdo do portal, conhecida como "Minha Página". Essa visão mostra apenas as categorias que interessam a cada usuário. O portal pode avisar ao usuário sempre que um novo conteúdo for adicionado às categorias por ele assinaladas. Os usuários podem publicar documentos no repositório corporativo para que esses sejam também visualizados por outros usuários. Essa geração privilegia a distribuição personalizada do conteúdo.
Terceira	Interativo	O portal incorpora aplicativos que melhoram a produtividade das pessoas e equipes, tais como correio eletrônico, calendários, agendas, fluxos de atividades, gerência de projeto, relatórios de despesas, viagens, indicadores de produtividade etc. Essa geração adiciona o caráter cooperativo ao portal, provendo múltiplos tipos de serviços interativos.
Quarta	Especializado	Portais baseados em funções profissionais, para gerência de atividades específicas na instituição, tais como vendas, finanças, recursos humanos, etc. Essa geração envolve a integração de aplicativos corporativos com o portal, de forma que os usuários possam executar transações, ler, gravar e atualizar os dados corporativos, e ainda incorpora outras possibilidades como comércio eletrônico, por exemplo.

Quadro 2: Gerações de portais corporativos

Fonte: Eckerson (1999 *apud* DIAS, 2001)

O portal corporativo evoluiu no sentido de disponibilizar conteúdos específicos para usuários definidos, procurando abrir um canal de compartilhamento entre os colaboradores e destes com os clientes da organização.

Atualmente, no cenário competitivo em que todas as organizações se encontram, arriscaríamos afirmar que a implantação de portais é fundamental em todas as organizações, pois informação e conhecimento são elementos essenciais para a gestão dos negócios. Cardoso (2002) aponta as seguintes vantagens dos portais:

- permitem aos departamentos de Tecnologia da Informação aumentar os benefícios dos *datawarehouses*⁶, *data marts*⁷ e outros projetos de *Business Intelligence*⁸;
- aumentam a eficiência da organização: os utilizadores pouparão tempo na busca da informação;
- aumentam a produtividade: melhor informação leva diretamente a melhores decisões. Adicionalmente, a possibilidade de tomar decisões rápidas significa freqüentemente diferença entre ganhar/perder um negócio;
- redução nos custos: o portal *self-service* reduz os custos, evitando que os administradores passem horas a produzir relatórios e documentos;
- *e-commerce*⁹: melhoramento nas relações com fornecedores, distribuidores e clientes;
- mais poderes para o utilizador final: o portal elevará a produtividade dos utilizadores, aumentando a sua independência.

Acrescentamos a essas vantagens o fato de possibilitar a integração, seja por meio de *chats*, fóruns, *blogs*, ou de qualquer outra ferramenta que permita o encontro e a troca de conhecimentos entre as pessoas. Dias (2003, p.69) declara que

[...] o grande desafio dos portais é o de ocultar os problemas relativos à heterogeneidade das fontes de informação e, assim, uma vasta produção científica tem se voltado para a integração de fontes heterogêneas de dados e a interoperabilidade de sistemas, com propostas que incluem a adoção de metadados, o uso de ontologias e outros recursos de integração [...]

Esses desafios encontram-se explicitados nas funcionalidades que os portais precisam apresentar, pois a heterogeneidade é inerente aos processos

⁶ “*Data warehouse* é um grande banco de dados que armazena dados de diversas fontes para futura geração de informações integradas, com base nos dados de funcionamento das funções empresariais operacionais de uma organização inteira” (REZENDE E ABREU, 2003, p.211).

⁷ *Data mart* se refere “a um depósito de dados que atende a certas áreas específicas da empresa e voltados para o processo decisório gerencial” (BARBIERI, 2001, p. 50).

⁸ Constitui-se na utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios das empresas.(BARBIERI, 2001)

⁹ Comércio eletrônico.

organizacionais. Sendo assim, deve-se encontrar uma forma de trabalhar com esses diferentes suportes, com vistas a proporcionar o compartilhamento das informações. Terra e Gordon (2002) dividem as funcionalidades em três grupos:

- Camada de Apresentação e Personalização define como os usuários enxergam e customizam a informação que é disponibilizada ou acessada pelo portal. As soluções mais intuitivas sendo oferecidas permitem aos usuários acessar a informação em ambientes mais relevantes e contextualizados.
- Taxonomias e Mecanismos de Busca determinam quão fácil será para os usuários encontrar informação relevante baseada em um conjunto de critérios de busca. [...] As soluções de busca têm-se desenvolvido para integrar variadas fontes de informação, tornando-se assim mais sofisticadas, intuitivas e adaptáveis às diferentes necessidades organizacionais.
- Aplicações Web. [...] Soluções sofisticadas de plataformas de portais fornecem uma série de recursos que tornam muito fácil integrar diferentes tipos de aplicações [...].

(TERRA E GORDON, 2002, p. 95)

Essas funcionalidades, representadas na Figura 2, sucintamente objetivam que o usuário encontre informações pertinentes ao desenvolvimento do trabalho e tenha a possibilidade de colaborar por meio da codificação e disponibilização de seus conhecimentos.

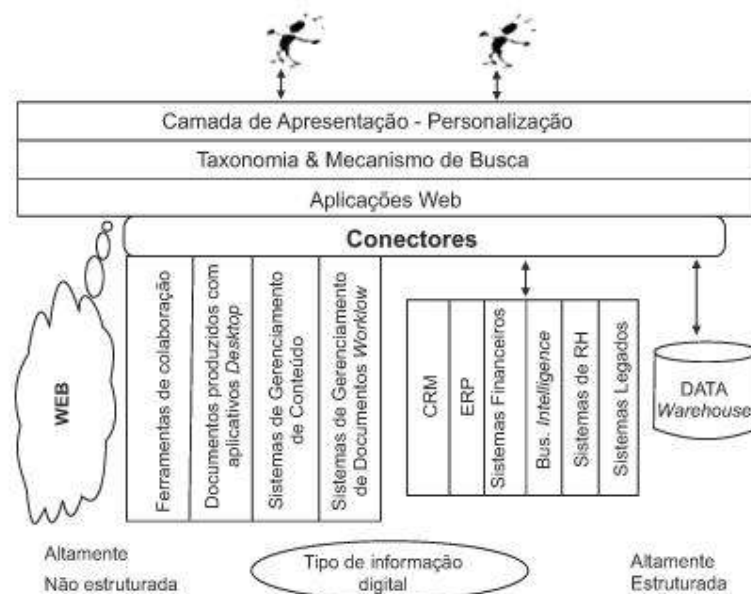


Figura 2: Componentes-chave da arquitetura PdCC

Fonte: Terra e Gordon (2002, p.96)

Os três grupos de funcionalidades - Camada de apresentação, Taxonomia e Mecanismo de busca e Aplicações Web - são suportados por conectores. Esses

conectores podem ser sistemas de gerenciamento de conteúdo, de documentos, ferramentas de colaboração, *softwares* como, CRM (*Customer Relationship Management*) que trabalha com dados integrados de clientes, BI (*Business Intelligence*), que dá suporte para o gerenciamento informacional da organização, visando à tomada de decisão, entre outros.

Analisando os portais sob a ótica da Gestão do Conhecimento, Terra e Gordon (2002) ressaltam três aspectos essenciais:

- encontrar informações relevantes e fontes de informação;
- codificar e publicar seus conhecimentos;
- colaborar on-line.

Esses aspectos estão intimamente relacionados às três camadas de funcionalidades apresentadas na Figura 2. Por outro lado, as soluções de busca, que resultam em encontrar informações relevantes e fontes de informação, remetem-nos à camada de organização, categorização e criação de taxonomias nesse ambiente, principalmente quando se trata de buscar informações não-estruturadas, fundamentais em um processo de Gestão do Conhecimento, pois delas derivam os conhecimentos tácitos.

Eckerson (1999 *apud* DIAS, 2001, p. 57) também apresenta 15 requisitos indispensáveis a um portal corporativo:

Requisito	Descrição
Fácil para usuários eventuais	Os usuários devem conseguir localizar e acessar facilmente a informação correta, com o mínimo de treinamento, não importando o local de armazenamento dessa informação. Encontrar informações de negócios no portal deve ser tão simples quanto usar um navegador Web. O portal deve servir como uma ponte para os usuários migrarem de métodos básicos de acesso e análise de informação para métodos mais sofisticados
Classificação e pesquisa intuitiva	O portal deve ser capaz de indexar e organizar as informações da empresa. Seu mecanismo de busca deve refinar e filtrar as informações, suportar palavras-chave e operadores booleanos, e apresentar o resultado da pesquisa em categorias de fácil compreensão. Tanto usando um mecanismo de busca ou navegando em uma estrutura de classificação, os usuários devem poder visualizar descrições coerentes de objetos antes de recuperá-los.
Compartilhamento colaborativo	O portal deve permitir aos usuários publicar, compartilhar e receber informações de outros usuários. O portal deve prover um meio de interação entre pessoas e grupos na organização. Na publicação, o usuário deve poder especificar quais usuários e grupos terão acesso a seus documentos/objetos.
Conectividade universal aos	O portal deve prover amplo acesso a todo e qualquer

recursos informacionais	recurso informacional, suportando conexão com sistemas heterogêneos, tais como correio eletrônico, bancos de dados relacionais e multidimensionais, sistemas de gestão de documentos, servidores Web, <i>groupware</i> , sistemas de áudio, vídeo etc. Para isso, deve ser capaz de gerenciar vários formatos de dados estruturados e não estruturados.
Acesso dinâmico aos recursos informacionais	Por meio de sistemas inteligentes, o portal deve permitir acesso dinâmico às informações nele armazenadas, fazendo com que os usuários sempre recebam informações atualizadas. O portal deve prover também acesso dinâmico a objetos criados por fornecedores de ferramentas de administração de documentos e <i>business intelligence</i> . Este acesso requer o desenvolvimento de interfaces de integração.
Roteamento inteligente	O portal deve ser capaz de direcionar automaticamente relatórios e documentos a usuários selecionados como parte de um processo bem definido de fluxo de informações.
Ferramenta de inteligência de negócios integrada	Para atender às necessidades de informação dos usuários, o portal deve integrar os aspectos de pesquisa, relatório e análise dos sistemas de inteligência de negócios.
Arquitetura baseada em servidor	Para suportar um grande número de usuários e grandes volumes de informações, serviços e sessões concorrentes, o portal deve basear-se em uma arquitetura cliente-servidor.
Serviços distribuídos	Para um melhor balanceamento da carga de processamento, o portal deve distribuir os serviços por vários computadores ou servidores.
Definição flexível das permissões de acesso	O administrador do portal deve ser capaz de definir permissões de acesso para usuários e grupos da empresa, por meio dos perfis de usuário. Estas permissões definem os tópicos ou categorias que os usuários têm acesso, canais que podem subscrever, funções que podem usar, os dados que podem ver e nível de interatividade com relatórios.
Interfaces externas	O portal deve ser capaz de se comunicar com outros aplicativos e sistemas, sendo possível ler e sincronizar outros serviços de diretório e interagir com os demais repositórios de informação.
Interfaces programáveis	O portal também deve ser capaz de ser "chamado" por outros aplicativos, tornando pública sua interface programável (API - <i>Application-Programming Interface</i>).
Segurança	Para salvaguardar as informações corporativas e prevenir acessos não autorizados, o portal deve suportar serviços de segurança, como criptografia, autenticação, <i>firewalls</i> etc. Deve também possibilitar auditoria dos acessos a informações, das alterações de configuração etc.
Fácil instalação e administração	Deve ser de fácil instalação, configuração e manutenção, e aproveitar, na medida do possível, a base instalada de <i>hardware</i> e <i>software</i> adquirida/contratada anteriormente pela organização. Deve ainda prover um meio de gerenciar todas as informações corporativas e monitorar o funcionamento do portal de forma centralizada e dinâmica.
Customização e personalização	O administrador do portal deve ser capaz de customizá-

lo de acordo com as políticas e expectativas da organização, assim como os próprios usuários devem ser capazes de personalizar sua interface para facilitar e agilizar o acesso às informações consideradas relevantes. Além disso, os usuários devem poder personalizar o portal, tornando-o mais fácil para seu próprio uso.

Quadro 3: Requisitos mínimos de um portal corporativo

Fonte: Eckerson (1999 *apud* DIAS, 2001)

Das quinze regras apresentadas por Eckerson, a da Classificação e pesquisa intuitiva, a da facilidade de localização da informação por usuários eventuais, a da autonomia para o usuário publicar e compartilhar seus conhecimentos e a da possibilidade de o usuário personalizar seu acesso estão relacionadas ao modelo de representação e organização da informação desenvolvido.

De acordo com Detlor (2000), os portais devem apresentar três componentes:

- esquema de classificação das informações em categorias;
- máquina de busca;
- *links* (para sites internos e externos).

Percebemos, nas funcionalidades e requisitos dos portais corporativos, que a representação e a organização da informação servem de base para os processos de integração e compartilhamento. Nessa perspectiva, Carvalho e Ferreira (2005) analisam portais, buscando características comuns com os Sistemas de Recuperação da Informação - (SRI) e, conseqüentemente, abrem caminhos para que os portais se beneficiem dos estudos da área. No Quadro 4, vemos um paralelo entre as funcionalidades de um portal e os módulos de um SRI. De acordo com Carvalho e Ferreira (2005), seus aspectos técnicos foram selecionados em Delphi Group (2000); Terra e Gordon (2002); Firestone (2003); Hazra (2003); Marcus e Watters (2002) e Portals Community (2003).

Aspecto Técnico do Portal	Função do SRI
Categorização	Representação (subsistema de entrada)
Mecanismo de Recuperação	Recuperação da Informação (subsistema de saída)
Gestão de Conteúdo	Seleção, Organização de Arquivos e Armazenamento (subsistema de entrada)
Apresentação/Personalização	Análise das Necessidades do Usuário (subsistema de saída)
Disseminação Seletiva	Disseminação (subsistema de saída)
Avaliação	Subsistema de Avaliação

Quadro 4: Comparativo das funcionalidades de um portal com um SRI

Fonte: Carvalho e Ferreira (2005, p.4)

Como observado no item 2.1, tanto o conhecimento tácito quanto o explícito encontram no portal corporativo meios de serem compartilhados e ampliados. O profissional da informação¹⁰ pode atuar como elo entre esses conhecimentos e seus usuários nesse ambiente. Em se tratando de conhecimento tácito, a possibilidade de contato entre os diversos especialistas de uma área, pelas chamadas ‘páginas amarelas’¹¹, seria um exemplo. Em relação ao conhecimento explícito, codificados e formalizados, os modelos de representação da informação oferecem os subsídios para uma recuperação eficiente e eficaz dos recursos existentes nas organizações (CHOO, 2000).

2.3 MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO

Modelos de representação do conhecimento são adotados na sistematização de conceitos de determinadas áreas com o objetivo de representar conteúdos informacionais para a recuperação. São exemplos as classificações, tesouros, ontologias e taxonomias.

Davis et al. (1992 *apud* CAMPOS, 2004, p.24) apresentam algumas características que nos auxiliam na compreensão do que são esses modelos:

- são mecanismos que propiciam um ‘pensar sobre o mundo’ em vez de interferir;
- representam o objeto, objetivando substituí-lo, carregando uma correspondência específica do objeto, mas de forma imprecisa;
- na escolha de uma representação em detrimento de outras, é feita também uma escolha de como e o que ver no ‘mundo’;
- essas representações permitem processos formalizados para se dizer ‘coisas sobre o mundo’, assim como, permitem a interpretação automática da informação.

O processo de representação é fundamental ao ser humano para a compreensão do mundo. Marcondes (2001, p.64) nos diz que “a vida real é uma

¹⁰ Visto aqui como aqueles que “disponibilizam informação em qualquer suporte; gerenciam unidades como bibliotecas, centros de documentação, centros de informação e correlatos, além de redes e sistemas de informação. Tratam tecnicamente e desenvolvem recursos informacionais; disseminam informação com o objetivo de facilitar o acesso e geração do conhecimento; desenvolvem estudos e pesquisas; realizam difusão cultural; desenvolvem ações educativas.”(CBO, 2002)

¹¹ Base de *links* não para o conhecimento, mas para os indivíduos que o detenham. (STEWART, 1998)

realidade demasiadamente vasta e bastante complexa para ser captada em uma totalidade pela mente humana”. É nesse momento que as correspondências entre os objetos, seu significado e seu significante se fazem necessárias. Em um processo cognitivo, o objeto recebe um substituto distinto dele, nas associações da mente humana.

De acordo com Campos (2004, p.24), os métodos de representação da informação e do conhecimento possibilitam, no âmbito da Ciência da Informação, “a elaboração de linguagens documentárias verbais e notacionais, visando à recuperação de informações e à organização dos conteúdos informacionais de documentos”. Em meio à crescente massa informacional disponibilizada pela Internet, os modelos de representação da informação e do conhecimento são indispensáveis para uma recuperação eficiente dessa informação. Marcondes (2001, p.67), referindo-se aos benefícios oriundos desses modelos, afirma que “a representação deve situar-se entre dois extremos para economizar energia e assim realizar seu papel: ser suficientemente rica sob o aspecto cognitivo e, ao mesmo tempo, sintética para economizar a energia do usuário de uma maneira significativa”. A riqueza cognitiva a que o autor se refere diz respeito à possibilidade da representação fazer com que o usuário, ao final da leitura ou da visualização, possa inferir na verdadeira pertinência da informação diante de suas necessidades.

Dodebei (2002) apresenta um ciclo informacional baseado em Lancaster, que trata a representação do conhecimento em seis etapas, conforme Figura 3. Essas etapas são divididas em dois subconjuntos: informação e documento.

O subconjunto informação constitui-se de forma independente, englobando a produção, registro e assimilação do conhecimento. Já o subconjunto documento, constituído de seleção e aquisição, organização da memória organizacional e disseminação da informação, necessita do subconjunto informação para existir. Na etapa de organização da memória organizacional, insere-se a representação do conhecimento.



Figura 3: Ciclo da Informação
Fonte: Dodebei (2002)

Para que o processo de organização e representação seja viabilizado, a construção de Linguagens Documentárias (LD) é imprescindível. Antes de abordarmos a definição de LD, é necessário entendermos o que é Linguagem Natural.

O ser humano constrói os significados da linguagem nas interações sociais, resultando no que chamamos de linguagem natural (LN). Segundo Cintra; et al. (2002, p.14), LN é aquela que “[...] analisa os dados da experiência segundo padrões que dependem da tradição cultural e do momento social do povo que a fala.” As linguagens documentárias surgem como uma forma de uniformizar um vocabulário específico a fim de tornar possível o intercâmbio e a recuperação da informação. De acordo com Cintra *et al.* (2002, p.33), “essas linguagens são construídas para indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos destinados a “traduzir” os conteúdos dos documentos.” Gardin *et al.* (1968 *apud* CINTRA *et al.*, 2002, p.35) definem a linguagem documentária como “um conjunto de termos, providos ou não de regras sintáticas, utilizadas para representar conteúdos de documentos técnico-científicos com fins de classificação ou busca retrospectiva de informações.” Esses autores prosseguem apresentando três elementos básicos que devem compor uma LD:

- um léxico, identificado como uma lista de elementos descritores, devidamente filtrados e depurados;
- uma rede paradigmática para traduzir certas relações essenciais e, geralmente estáveis, entre descritores;
- uma rede sintagmática destinada a expressar as relações contingentes entre os descritores, relações que são válidas no contexto particular onde aparecem. (GARDIN *et al.*, 1968 *apud* CINTRA *et al.*, 2002, p.35).

Léxico é definido como “o conjunto de unidades reais e virtuais que formam a língua de uma comunidade [...]” (CINTRA *et al.*, 2002, p.36). Unidades virtuais são as possibilidades potenciais da língua que permitem a atividade humana da fala. Nesse sentido, um vocabulário controlado seria apenas um léxico que, segundo Lancaster (1987), poderia ou não apresentar uma rede de relações essenciais. A LD vai além e busca estabelecer relações significativas entre os termos, procurando entender o contexto para compreender o significado.

As práticas de elaboração de modelos de representação do conhecimento analisados nesta pesquisa – taxonomias - dependem das características e objetivos do ambiente em que esses modelos serão desenvolvidos; aqui são direcionados ao ambiente de portais corporativos.

Para melhor entender as linguagens documentárias, são descritos a seguir os tipos de modelos de representação do conhecimento.

2.3.1 CLASSIFICAÇÃO

O ato de classificar é intrínseco aos processos mentais humanos. Quando separamos ‘coisas’ em grupos distintos com características comuns entre si, estamos classificando. Podemos classificar conceitos de forma objetiva, como, por exemplo, grupo de peças circulares e grupo de peças quadradas, ou de uma forma mais subjetiva, bom ou mau. De acordo com Dahlberg (1978a), a classificação está buscando, “ao invés da apresentação sistemática das classes”, os “conceitos orientados para o objeto”. Conceito, segundo a definição de Dahlberg (1978a), “é a unidade de conhecimento que surge pela síntese dos predicados necessários relacionados com determinado objeto e que, por meio de sinais lingüísticos, pode ser comunicado”.

Merriell (1958, *apud* PIEDADE, 1977), sob a ótica da Biblioteconomia, define ‘classificar’ como “a arte de dar aos livros um lugar exato num sistema de

classificação, no qual os vários ramos do saber ou a descrição da vida humana, em seus vários aspectos, estão agrupados conforme suas semelhanças ou suas relações recíprocas”. Em se tratando de materiais bibliográficos como livros, em que a representação da informação precisa ser unívoca, os sistemas de classificação tradicionais continuam sendo as formas mais viáveis. Por isso, nos sistemas ‘tradicionais’ de classificação, a ordenação por classes ainda predomina, os materiais são ‘encaixados’ dentro de classes pré-definidas.

Um sistema de classificação é caracterizado por Piedade (1977, p.9) como sendo “um conjunto de classes apresentadas em ordem sistemática. É uma distribuição de um conjunto de idéias por um certo número de conjuntos parciais, coordenados e subordinados”. Para Langridge (1977), um sistema de classificação é “um mapa completo de qualquer área do conhecimento, mostrando todos os seus conceitos e suas relações”. Essa forma de representar o conhecimento varia conforme o sistema de classificação e representa sempre uma forma de conceber o mundo, uma visão científica, política, pedagógica ou uma combinação delas.

Piedade (1977) aponta a existência de dois sistemas distintos de classificação, o filosófico e o bibliográfico. Enquanto os sistemas filosóficos apresentam a “[...] finalidade de definir, esquematizar e hierarquizar o conhecimento, preocupados com a ordem das ciências ou a ordem das coisas”, as classificações bibliográficas “[...] servem de base à organização de documentos nas estantes, em catálogos, bibliografias, etc.” (PIEADADE, 1977, p.52). Svenonius (1985 *apud* ARAÚJO, 2006) também identifica essas duas formas de classificação, uma preocupada com a classificação do conhecimento propriamente dita, de forma filosófica e a outra com a pretensão de tornar acessíveis (no que tange à organização e recuperação) esses conhecimentos. Neste trabalho, ocupar-nos-emos apenas das classificações bibliográficas.

As classificações bibliográficas objetivam facilitar a localização do documento; porém, todas trazem uma visão de mundo peculiar que marca as concepções políticas e culturais nas quais estão inseridas. De acordo com Piedade (1977), as classificações bibliográficas têm como finalidades:

- organizar os documentos nas estantes e/ou arquivos;
- organizar as referências bibliográficas e/ou fichas em catálogos.

Na área biblioteconômica, os sistemas de classificações bibliográficas mais conhecidos e utilizados são a Classificação Decimal de Dewey (CDD), a

Classificação Decimal Universal (CDU), a Classificação Facetada e a Classificação da Biblioteca do Congresso dos Estados Unidos (LCC – Library of Congress Classification), que são codificadas por notações, e serão abordadas a seguir.

a) Classificação Decimal de Dewey (CDD)

Este sistema de classificação tem como autor Mevil Dewey. A CDD foi desenvolvida com base no sistema de classificação de Harris, essência das idéias de Bacon, que via o homem como o centro do universo (LIMA, 2004). Teve sua primeira edição publicada em 1876 e procurava por meio de suas classes principais abarcar todo o conhecimento produzido.

É composta por classes gerais, subclasses e tabelas auxiliares; sua abordagem vai do conhecimento geral para o mais específico, como demonstrado no exemplo da área de Ciências Naturais e Matemáticas (LIMA, 2004):

000 Generalidades

100 Filosofia e Psicologia

200 Religião

300 Ciências Sociais

400 Linguagem

500 Ciências Naturais e Matemáticas

510 Matemática

520 Astronomia

530 Física

540 Química

541 Química física e teórica

542 Técnicas. Equipamentos e Materiais

543 Química analítica

544 Análise qualitativa

550 Ciências da Terra

560 Paleontologia

570 Biologia

580 Botânica

590 Zoologia

600 Tecnologia (Ciências Aplicadas)

700 Artes

800 Literatura e Retórica

900 Geografia e História

Como podemos notar no exemplo, a CDD utiliza números decimais de forma hierárquica. A CDD contém as áreas fundamentais do conhecimento e não apresenta um princípio explícito (LANGRIDGE, 1977). Esse sistema de classificação está em constante atualização e, em sua 22ª edição, continua a ser muito utilizado

em bibliotecas no mundo todo. O que, segundo Langridge (1977), não se justifica, devido à existência de sistemas mais aprimorados, com um caráter de análise e síntese das informações.

b) Classificação Decimal Universal (CDU)

A CDU foi idealizada por Paul Otlet e Henri de La Fontaine. Originou-se da tentativa de sistematizar e organizar a produção científica. Para tanto, seus autores contaram com o apoio do Instituto Internacional de Bibliografia, que originou a Federação Internacional de Informação e Documentação (FID). Com base na Classificação Decimal de Dewey (CDD), expandiram e modificaram as classes com o objetivo de criar uma classificação mais detalhada. De acordo com Silva e Ganim (1994), baseando-se nas características da CDD,

[...] procuraram incorporar-lhe uma outra característica que enriqueceria profundamente o sistema original [...], a introdução do conceito de relação e dos recursos para representá-la através da síntese, que permite a formação de notações compostas para representar conceitos novos não previstos no sistema. (SILVA e GANIM, 1994, p.06)

Sua primeira edição em francês é de 1905. As dez classes principais da CDU assemelham-se às da CDD; a classe quatro foi transportada para a oito na década de 1960, encontrando-se vaga no momento. São elas:

- 0 Generalidades**
- 1 Filosofia**
- 2 Religião. Teologia**
- 3 Ciências Sociais**
- 4 Classe Vaga**
- 5 Ciências Puras**
- 6 Ciências Aplicadas**
- 7 Belas Artes. Divertimentos. Desportos**
- 8 Filologia e Literatura**
- 9 Geografia. Biografia. História**

Silva e Ganin (1994) citam como características fundamentais da CDU:

- Decimalidade: divisão do conhecimento em dez grandes classes e a subdivisão destas em quantas forem necessárias para atingir a especificidade requerida.

- Universalidade: pretensão e capacidade para representar todo o conhecimento universal, prevendo crescimento. Seus símbolos são de conhecimento e emprego unívocos.
- Caráter hierárquico: classificação estruturada em partes subordinadas, do geral ao mais específico.

Langridge (1977) acredita que a característica da universalidade é discutível e diz respeito apenas “a cobertura de assunto, uma vez que nunca foram estabelecidas regras para uso internacional”. Até mesmo pela grande flexibilidade do sistema, duas classes principais podem se fundir para criar um único número, dependendo das necessidades. A idéia de Otlet e La Fontaine era a criação de uma classificação que levasse em conta as várias particularidades da documentação, sendo para isso muito flexível.

c) Classificação da Biblioteca do Congresso dos Estados Unidos (LCC – Library of Congress Classification)

A classificação da Biblioteca do Congresso dos Estados Unidos surgiu de necessidades específicas dessa instituição. É baseada na já extinta Classificação Expansiva de Cutter (LANGRIDGE, 1977). Piedade (1977) diz que, no momento da criação do sistema, “a extensão de cada classe, isto é, sua maior ou menor especificidade dependeu do acervo da biblioteca do Congresso”; porém, com a adoção do sistema por instituições do mundo todo, a classificação foi expandida e adaptada aos diferentes contextos.

As classes gerais são representadas por letras do alfabeto, seguidas de números arábicos nas subdivisões, conforme exemplo da classe Q, abaixo (LIMA, 2004):

Q Ciência

QA Matemática

QB Astronomia

QC Física

QD Química

QD 1-65 Generalidades incluindo Alquimia

QD 71-142 Química analítica

QD 146-197 Química inorgânica

No seu desenvolvimento, cada classe ficou sob a responsabilidade de um grupo de trabalho. Desde então, cada classe é publicada em volumes independentes. As tabelas auxiliares também foram desenvolvidas de acordo com a necessidade da classe; entretanto, algumas delas não possuem essas tabelas.

Por ter sido concebido a partir de uma necessidade prática, esse sistema sofre críticas por não dispor de um campo teórico para embasá-lo. Não apresenta flexibilidade e facilidade na determinação dos assuntos. Suas atualizações dependem das equipes responsáveis, levando a uma lentidão que dificulta o acompanhamento do crescimento do conhecimento.

d) Classificação Facetada

Encontrada na literatura também sob a denominação 'Classificação de Dois Pontos' (LIMA, 2004; LANGRIDGE, 1977), este sistema de classificação foi uma grande contribuição de Shiyali Ramamrita Ranganathan, matemático e bibliotecário indiano.

Pela primeira vez, foi desenvolvida juntamente com a tabela uma teoria que a embasava, fazendo com que a classificação ultrapassasse a barreira de técnica e assumisse uma construção teórica com suas reflexões e problemas pertinentes. A classificação facetada teve sua primeira edição, a *Colon Classification*, em 1933.

De acordo com Araújo (2006), Ranganathan, com o desenvolvimento deste novo sistema de classificação, buscava maior flexibilidade “que permitisse a síntese, o agrupamento de vários componentes para a especificação do assunto”. Este sistema de classificação procura promover relações variadas entre os conceitos. Para a sua elaboração, foi necessário determinar quais conceitos e relações iriam ser contemplados. De acordo com Campos (2001), na Classificação Facetada “[...] as unidades classificatórias são o assunto básico e a idéia isolada”. Assunto básico é definido como um “assunto sem nenhuma idéia isolada como componente”; idéia isolada, como “alguma idéia ou complexo de idéias ajustadas para formar um componente de um assunto, mas, em si mesma, ela não é considerada um assunto”

(RANGANATHAN, 1967 *apud* CAMPOS, 2001, p.49). Podemos citar como exemplo o assunto 'Fabricação de cadeiras', em que 'cadeira' seria a idéia isolada que necessita do complemento 'fabricação' para tornar-se um assunto. Isso caracteriza a idéia subjacente à classificação facetada que prioriza a flexibilidade e a necessidade de combinação dos termos.

Segundo Barbosa (1972 *apud* LIMA, 2004, p.59), "uma classificação facetada é um sistema que agrupa termos estruturados, na base da análise de um assunto, para identificação de suas facetas", que são "[...] combináveis no ato de classificar, para produzir devidamente o tema dos documentos" (PIEIDADE 1977). Barbosa (1969 *apud* ARAÚJO, 2006) afirma ainda que, "segundo Ranganathan, analisar um assunto por facetas significa que cada aspecto desse assunto pode ser visto como as manifestações de certas características ou facetas que obedecem a postulados pré-determinados". Sendo assim, essa análise pode apresentar resultados ilimitados, tantos quantos forem as características encontradas.

As facetas são manifestações das categorias fundamentais. As categorias fundamentais demarcam a primeira classificação de assuntos dentro de um grande universo. Para Campos (2001), "são elas que fornecem a visão de conjunto dos agrupamentos que ocorrem na estrutura, possibilitando assim o entendimento global da área". Langridge (1977) também entende as categorias fundamentais como uma estrutura mais abrangente da classificação e as facetas como manifestações das categorias em diferentes classes. Ranganathan (1962) instituiu cinco categorias fundamentais:

Personalidade – É considerada indefinível, quando um termo não se adapta a nenhuma outra categoria, é entendida como uma manifestação desta.

Matéria – Materiais constitutivos de um objeto. Ex.: mesa/Matéria: madeira.

Energia – Processos, operações, ações, técnicas, métodos, fenômenos.

Espaço – Lugares.

Tempo – Períodos.

O grande mérito deste sistema foi a contemplação das relações entre os termos, formando um todo coerente. Pela organização hierárquica, flexível e pós-coordenada, é possível visualizar o conceito e suas possíveis relações, que variam para cada documento. Esse pode ser considerado o primeiro passo em direção à construção de tesauros.

2.3.2 TESAUROS

O termo 'tesauro' é de origem greco-latina, *thesaurus*, que significa tesouro. De acordo com Cavalcanti (1978), as obras que introduziram o termo foram o *Thesaurus Linguae Romanae et Britannica*, em 1565, de Cooper e o *Thesaurus of English Words and Phrases de Roget*, em 1852. O tesauro, na obra de Roget, era entendido como um dicionário em que, em vez de conter uma lista de palavras com o respectivo significado, continha significados com possíveis palavras para expressá-los. Esse conceito de tesauro foi 'importado' pelas áreas que trabalham com informação. Dodebei (2002) explica que foi utilizado na Ciência da Informação, principalmente na recuperação da informação, "[...] como sendo um instrumento capaz de transportar conceitos e suas relações mútuas, tal como expressos na linguagem dos documentos, em uma língua regular, com controle de sinônimos e estruturas sintáticas simplificadas"; passou então a constituir-se em uma lista de termos, um vocabulário controlado para a indexação. Para Cavalcanti (1978, p.30), tesauro "é a lista estruturada de termos associados, empregados por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entrada, e para permitir aos pesquisadores a recuperação da informação que procuram". Essa definição está relacionada ao aspecto funcional estabelecido pelo programa Unisist na norma ISO 2788-1974, que conceitua um tesauro como sendo "um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários, numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita"¹². Sob o aspecto estrutural, o tesauro é entendido como "um vocabulário controlado e dinâmico de termos relacionados semanticamente e genericamente que cobre um domínio específico do conhecimento"¹³ (ISO 2788 – 1974). O aspecto funcional está mais direcionado ao documento e à sua recuperação pelo usuário; já o aspecto estrutural focaliza principalmente a representação de áreas específicas do conhecimento.

¹² "[...]is a terminological control device used in translating from the natural language of documents, indexers, or users into a more constrained "system language" (documentation language, information language)" (ISO 2788-1974, p.121).

¹³ "[...]is a controlled and dynamic vocabulary of semantically and generically related terms which covers a specific domain of knowledge" (ISO 2788-1974, p.121).

A ANSI/NISO Z39-19-1993, norma internacional para a elaboração de tesouros publicada pela *National Standards Organization* (NATIONAL INFORMATION STANDARDS ORGANIZATION), une esses dois tipos. No que diz respeito ao aspecto estrutural, a norma define tesouro como sendo “um vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, de homografias, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padronizados, que devem ser empregados de forma recíproca”¹⁴. Conforme Cavalcanti (1978), esses indicadores de relações podem ser:

Termo genérico (TG), termo específico (TE) e termo relacionado (TR)

Fruta

TG1 Produto agrícola

TE1 Fruta cítrica

TE2 Laranja

TE2 limão

TE2 Lima

TR Agricultura

Relação de Equivalência

Tangerina use Mexerica

Mexerica UP Tangerina

Homógrafos

Laranja (fruta)

Laranja (cor)

Quanto ao aspecto funcional, a norma ANSI-NISO Z39.19-2003 define quatro funções para tesouros:

Tradução: para prover um modo de traduzir a linguagem natural dos autores, indexadores e usuários em um vocabulário controlado usado para indexação e recuperação.

Consistência: para promover consistência na designação de termos de indexação.

¹⁴ “A controlled vocabulary arranged in a known order in which equivalence, homographic, hierarchical, and associative relationships among terms are clearly displayed and identified by standardized relationship indicators, which must be employed reciprocally” (NATIONAL INFORMATION STANDARDS ORGANIZATION - ANSI-NISO Z39.19-2003, p.1).

Indicação de relacionamentos: para indicar relacionamentos semânticos entre termos.
Recuperação:¹⁵ para servir como auxílio de busca na recuperação de documentos.

O tesauro auxilia tanto o usuário na recuperação da informação quanto o indexador no momento de inserir um documento no sistema. Assim, como em outros modelos de representação do conhecimento, questões de sinonímia, homógrafos e polissemias devem ser tratadas, objetivando a qualificação dos processos.

O termo tesauro documentário surgiu da necessidade de se tratar grandes quantidades de informação, com linguagem mais específica e direcionada (DODEBEI, 2002); quantidade essa que teve o crescimento maximizado com a Internet e as novas tecnologias da informação e comunicação. O tesauro, estando relacionado ao processo de recuperação da informação e à automação dos sistemas, sofreu algumas mudanças, ganhando mais rapidez e eficiência na tradução dos termos e nas suas relações pós-coordenadas¹⁶. Ainda hoje, “o tesauro continua a ser o instrumento de organização do pensamento que auxilia o usuário a identificar o termo mais adequado à representação de sua necessidade de informação” (GOMES, 1996, p.56), pois a precisão dos termos de indexação é essencial no momento da recuperação.

2.3.3 ONTOLOGIA

A palavra ontologia deriva do grego *onto* (ser) + *logia* (estudo), ciência que estuda o ser, como tal. As primeiras definições de ontologia surgiram na Filosofia, com forma generalista, no século XVII. O filósofo Husserl definia ontologia como sendo a ciência das essências, dividindo-as em:

Formal – fundamento de todas as ciências e interessando-se pelas essências e,

Material – conjunto de ontologias setoriais que se preocupa com os fatos.

¹⁵ “Translation: To provide a means for translating the natural language of authors, indexers, and users into a controlled vocabulary used for indexing and retrieval.

Consistency: To promote consistency in the assignment of index terms.

Indication of Relationships: To indicate semantic relationships among terms.

Retrieval: To serve as a searching aid in retrieval of documents.” (NATIONAL INFORMATION STANDARDS ORGANIZATION - ANSI-NISO Z39.19-2003, p.1).

¹⁶ As relações pós-coordenadas dos termos são aquelas que “[...] utilizam conceitos simples na indexação, e a combinação ou coordenação de conceitos para obtenção de conceitos compostos é feita no momento da recuperação”.(PIEDADE, 1977, p.2)

O conceito de ontologia, no século XVII, estava ligado à metafísica, designando 'aquilo que existe'.

Na representação do conhecimento, conceito da Ciência da Informação, o sentido de ontologia é diferente. Com base em uma determinada área de domínio, a ontologia se propõe a classificar as coisas em categorias, na perspectiva do sujeito e da linguagem do domínio. Partindo da definição de Gruber (1996, p.1), em que “uma ontologia é uma especificação explícita de uma conceitualização”¹⁷, Guarino e Giaretta (1995 *apud* ALMEIDA e BAX, 2003) dizem que “o principal problema com tal interpretação é que ela é baseada na noção de conceitualização, a qual não corresponde à nossa intuição [...]”.

Uma conceitualização é um grupo de relação extensional descrevendo um 'estado das coisas' particular, enquanto a noção que temos em mente é uma relação intensional, nomeando algo como uma rede conceitual a qual se superpõe a vários possíveis 'estados das coisas'. Nesse sentido, Dahlberg (1978b), em seu artigo sobre a teoria do conceito, trata da intensão e extensão, dizendo que, “a intensão do conceito é a soma total das suas características [...]”, sejam elas genéricas ou específicas, sendo um contido no outro. Já a extensão do conceito é “entendida como a soma total dos conceitos mais específicos”, diretamente relacionados com o termo.

Sob o aspecto intensional, Guarino (1998 *apud* ALMEIDA e BAX, 2003) afirma que “[...] ontologia se refere a um artefato constituído por um vocabulário usado para descrever uma certa realidade, mais um conjunto de fatos explícitos e aceitos que dizem respeito ao sentido pretendido para as palavras do vocabulário”. Ou seja, uma ontologia é construída com base nas relações intensionais, usadas involuntariamente pelo ser humano, formando assim redes relacionais. Segundo Neches (1991 *apud* FEITOSA, 2005), “uma ontologia define os termos básicos e as relações, compreendendo o vocabulário de uma área de tópico, bem como as regras para a combinação de termos e as relações para definir as extensões do vocabulário”. Desenvolvida de forma consensual e formal, “uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada” (BORST, 1997, p.12 *apud* ALMEIDA e BAX, 2003). Almeida e Bax (2003, p.9) classificam essa definição como simples e completa, afirmando que

¹⁷ “An ontology is an explicit specification of a conceptualization”. (GRUBER, 1996)

nessa definição, “formal” significa legível para computadores; “especificação explícita” diz respeito a conceitos, propriedades, relações, funções restrições, axiomas, explicitamente definidos; “compartilhado” quer dizer conhecimento consensual; e “conceitualização” diz respeito a um modelo abstrato de algum fenômeno do mundo real.

As ontologias são inerentes aos estudos de Web Semântica pelo ‘caráter’ de metadados que apresentam, objetivando o processamento automatizado da informação. Gruber (1996); Noy e Guinness (2001) citam como componentes básicos de uma ontologia:

- classes (organizadas em uma taxonomia);
- relações (representam o tipo de interação entre os conceitos de um domínio);
- axiomas (usados para modelar sentenças sempre verdadeiras);
- instâncias (utilizadas para representar elementos específicos, ou seja, os próprios dados).

Ontologias aplicadas em portais de informação, segundo Casanova (2003), “uniformizam a descrição do conteúdo fornecido e indexado por membros das comunidades, ou por provedores de conteúdo na Web [...]”, sendo esse procedimento fundamental para o compartilhamento das informações e conhecimentos desenvolvidos no âmbito organizacional. As ontologias organizam o conhecimento em forma de uma teia de relações, assim como o cérebro humano, em uma relação intensional. Essa forma de organização corrobora a afirmação de Almeida e Bax (2003, p.17) de que as ontologias “[...] permitem formas de representação baseadas em lógica, o que possibilita o uso de mecanismos de inferência para criar novo conhecimento a partir do existente”. Muito utilizada em processos de Gestão do Conhecimento em que é necessário organizar o conhecimento tácito dos sujeitos, a ontologia possibilita a formação de relações entre conceitos, no nível da semântica e propicia a representação de várias relações de um mesmo objeto em poucas definições.

2.3.4 TAXONOMIA

A explosão informacional, decorrente dos avanços tecnológicos, da crescente robotização e automação das empresas, faz-nos antever profundas modificações nos ambientes corporativos. Diante desse quadro e motivada pela evolução

tecnológica, em que os usuários adquirem um papel essencial tanto na produção como na categorização e uso das informações, surgiu a necessidade de meios automatizados de pesquisa e filtragem de informação, com o uso de taxonomias.

O termo taxonomia vem do grego *taxis*=ordem e *onoma*=nome e derivou-se de um dos ramos da Biologia que trata da classificação lógica e científica dos seres vivos, fruto do trabalho do médico e botânico sueco Carolus Linnaeus (ou Karl von Linné). Porém, apesar de as taxonomias serem derivadas da Biologia, nos ambientes digitais seu aparecimento e uso, segundo autores como Edols (2001), Adams (2000) e Plosker (2005), estão relacionados com as formas automatizadas de criação da informação, tornando-se alvo de estudos da Ciência da Informação.

No âmbito da Gestão do Conhecimento, as taxonomias são definidas como “[...] elementos estruturantes, estratégicos e centrais para negócios baseados em informação e conhecimento [...] para classificar e facilitar o acesso à informação” (TERRA, GORDON, 2002).

Para Martinez et al. (2004, p.106), “a taxonomia, em um sentido amplo, é a criação da estrutura (ordem) e dos rótulos (nomes) que ajudam a localizar a informação relevante. Em um sentido mais específico, é o ordenamento e rotulação de metadados, que permite organizar sistematicamente a informação primária”. É interessante notar que a estrutura mais citada na literatura para a taxonomia é a hierárquica, sendo esta uma forma de caracterizá-la, como podemos perceber na Figura 4:

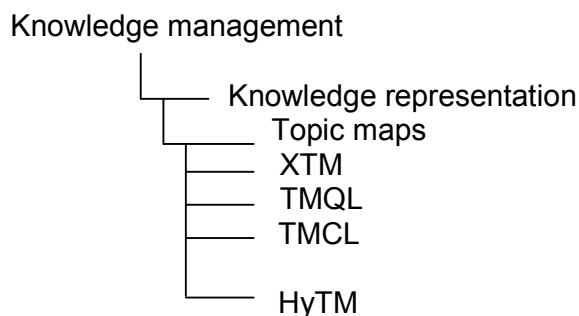


Figura 4: Exemplo de taxonomia
Fonte: Garshol (2004)

A taxonomia organiza a informação da mais genérica à mais específica, utilizando-se da relação gênero-espécie entre os termos. Esse tipo de relação é definido por Dahlberg (1978b) como sendo o que aparece “entre dois conceitos que

têm idênticas características, sendo, porém, que de uma relação à outra é apresentada uma característica adicional, de modo que surge entre eles uma hierarquia”. No entanto, as taxonomias corporativas não se limitam a essa relação somente, como veremos adiante.

O objetivo da taxonomia não é unicamente classificar e facilitar o acesso à informação, mas, igualmente,

[...] representar conceitos através de termos, agilizar a comunicação entre especialistas e outros públicos; encontrar o consenso; propor formas de controle da diversidade de significação e oferecer um mapa de área que servirá como guia em processo de conhecimento. (TERRA, et al., 2005, p.1)

Adams (2000) explica que “[...] taxonomias organizam e criam significados para os relacionamentos entre coisas ou idéias”¹⁸. Essa definição faz com que autores como Almeida (2005), por exemplo, vejam as ontologias como extensões taxonômicas ou, ainda, taxonomias como modelos de tesouros.

As organizações encontram dificuldade em prover acesso a informações semi-estruturadas (que podem estar nos servidores da empresa ou com grupos de usuários), e essa dificuldade é muito maior quando tratamos de informações que não apresentam estrutura alguma (armazenadas em arquivos pessoais, de acordo com a necessidade dos indivíduos), mas que podem ser importantes para outras pessoas dentro da organização e precisam ser compartilhadas. As taxonomias promovem o elo entre os indivíduos e a informação.

Taxonomias estão voltadas para a organização das informações em ambientes específicos, visando à recuperação eficaz. Para isso, estabelecem parâmetros em todo o ciclo de produção informacional¹⁹, no qual profissionais distribuídos por espaços físicos distintos participam do processo de criação do conhecimento de forma organizada. Dutra e Busch (2003) apresentam uma pesquisa realizada no *Yahoo* (Figura 4) com o termo ‘*trial*’, antes e depois do desenvolvimento de uma taxonomia. Percebe-se que a taxonomia instaura uma organização lógica e interdependente entre os termos, facilitando tanto a busca quanto o conhecimento de uma determinada área do conhecimento.

¹⁸ “[...]Taxonomies order and give meaning to the relationship among things or ideas”.(ADAMS, 2000, p.2)

¹⁹ Abordado no item 2.3, p.32.

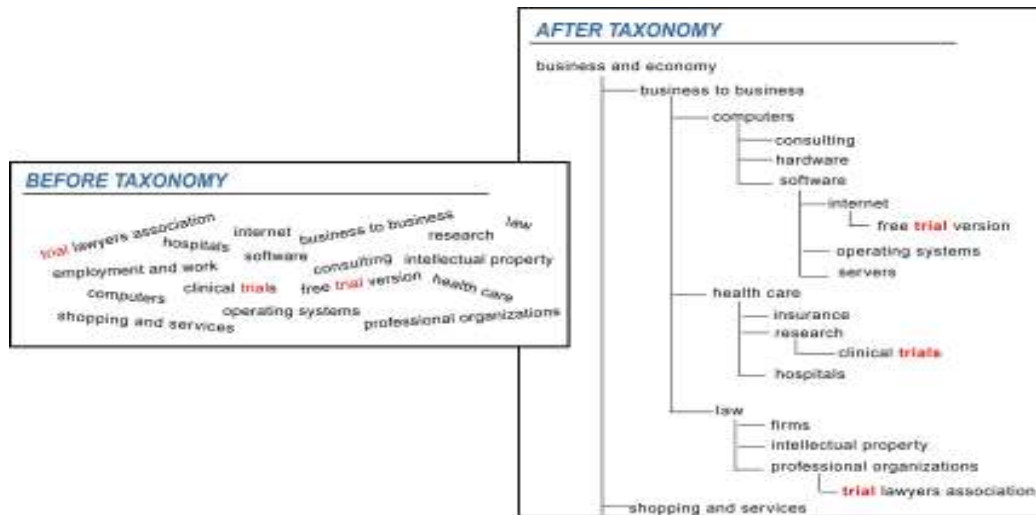


Figura 5: Pesquisa com o termo 'trial'
Fonte: Dutra e Busch (2003)

Como tratado anteriormente, o conceito de taxonomia algumas vezes é limitado a termos organizados em categorias hierárquicas; entretanto, quando tratamos de taxonomias em ambientes corporativos, já encontramos definições mais específicas.

Conway e Sligar (2002) classificam as estruturas taxonômicas, aplicadas em ambientes corporativos, em três tipos, de acordo com o objetivo a ser alcançado: taxonomia descritiva, de navegação e taxonomia para gerenciamento de dados. A taxonomia descritiva consiste na criação de um vocabulário controlado com o objetivo de otimizar a busca e recuperação das informações. Baseia-se na estrutura de um tesouro, com o estabelecimento dos termos significativos em um determinado contexto, trabalhando com as relações de sinonímia, homonímia, etc. O objetivo é fazer com que essa taxonomia tenha significado no ambiente em que foi construída, buscando a padronização da linguagem a fim de que uma comunicação efetiva aconteça.

A taxonomia de navegação é baseada nos modelos mentais de organização das informações. A estruturação das informações fica visível ao usuário final, pois o objetivo é que a descoberta das informações aconteça na navegação. E, com base na observação desse comportamento de busca, a taxonomia vai modificando-se. Conhecer o usuário nesse tipo de taxonomia é fundamental. As relações entre os termos precisam fazer sentido para os usuários dessa taxonomia, sem necessariamente apresentarem as subordinações lógicas.

A taxonomia para gerenciamento de dados tem como objetivo facilitar o compartilhamento em grupos específicos da organização, por exemplo, o setor de vendas. É composta de uma pequena lista de termos autorizados sem qualquer estrutura hierárquica, assemelha-se com as taxonomias descritivas; não tem, todavia, a pretensão de representar e promover acesso a toda informação produzida na organização.

Quanto à estruturação de uma taxonomia corporativa, Blackburn (2006) parte do pressuposto de que usualmente elas são hierárquicas e as divide em três tipos básicos: por assunto, por unidade de negócio e funcional.

A taxonomia por assunto, segundo a autora, faz uso de um vocabulário controlado. Organiza os termos por assuntos, partindo do mais geral aos mais específicos, por ordem alfabética, exigindo do usuário o conhecimento da área.

A taxonomia por unidade estabelece sua organização seguindo as unidades de negócio da organização. Apresenta como vantagem o fato de os indivíduos já estarem familiarizados com essa estrutura, não apresentando dificuldades em adotá-la. Entretanto, sempre que houver mudanças nessa estrutura, a taxonomia obrigatoriamente será modificada. Blackburn (2006) também mostra a dificuldade em trabalhar com documentos gerenciados ou compartilhados por várias unidades de negócios.

A taxonomia funcional é organizada levando-se em consideração as funções e atividades desenvolvidas pela organização. Os processos organizacionais mais gerais são utilizados para determinar também os níveis mais gerais da taxonomia. Já as funções formam os níveis intermediários, e as atividades o nível mais baixo, em uma estrutura hierárquica. Por exemplo, em uma agência reguladora do governo, Licenças e Certificações seriam o processo geral; Inspeção, a Função; Notificação, a atividade. Apesar da divisão em três tipos distintos, Blackburn (2006) indica como mais apropriada a combinação das estruturas, levando em consideração o ambiente em que serão desenvolvidas.

Verifica-se que as classificações dos tipos de taxonomia propostas coincidem em alguns pontos. A taxonomia descritiva (CONWAY; SLIGAR, 2002) e a taxonomia por assunto (BLACKBURN, 2006) referem-se ao uso de um vocabulário controlado que estabelece a organização e controle de significação dos termos, objetivando a otimização da busca. Assim como a taxonomia de gerenciamento de dados (CONWAY; SLIGAR, 2002) e a taxonomia funcional (BLACKBURN, 2006) visam

representar a informação de setores específicos da organização. Podemos perceber que uma taxonomia corporativa, como forma de representar toda a informação disponível na organização (WOODS, 2004), não pode se restringir a apenas uma tipologia. A taxonomia corporativa precisa combinar os diferentes tipos a fim de desenvolver uma co-relação das diferentes linguagens funcionais usadas na organização (por meio da padronização dos termos e construção das relações semânticas); fornecer suporte para o mecanismo de busca (pela categorização dos resultados de busca); fornecer suporte e ferramentas de auxílio a um mecanismo de navegação (com base na estrutura taxonômica); além de conferir autoridade para qualquer usuário nomear e compartilhar documentos e outras informações no portal, buscando a explicitação do conhecimento (GILCHRIST, 2001).

Segundo Woods (2004), “uma taxonomia clássica supõe que cada elemento só pode pertencer a um ramo da árvore hierárquica. No entanto, em um ambiente corporativo isso não é possível nem desejável” (WOODS, 2004)²⁰. Um documento pode ser de interesse de vários departamentos dentro da organização, com implicações e objetivos diferentes e precisa estar representado dentro desses diversos interesses, sem necessariamente se constituir apenas em uma relação gênero-espécie.

Outros tipos de relação entre conceitos, como as de funcionalidade também são importantes. A relação funcional é aquela em que, segundo Dahlberg (1978a) “pode-se conhecer o caráter semântico [...] tendo por base as chamadas valências semânticas dos verbos [...]”. Valência semântica é caracterizada como “a soma dos lugares a serem preenchidos de acordo com a ligação deste conceito com outros” (DALHBERG, 1978a). As relações entre os conceitos também podem ser compostas pelas relações intensionais, redes de conceitos, visto nas ontologias e essenciais nas taxonomias. A crescente produção informacional nas organizações, associada à disponibilização pela Internet de um volume cada vez maior de dados, faz com que a manutenção do controle sobre essa produção se torne uma tarefa praticamente impossível. Dessa maneira, a participação de representantes de todos os segmentos da organização no desenvolvimento da taxonomia permite ao usuário desempenhar

²⁰ “A classical taxonomy assumes that each element can only belong to one branch of the hierarchical tree. However, in a corporate environment, such formal ordering is neither feasible nor desirable”. (WOODS, 2004)

o papel de criador, revisor, editor e publicador dos documentos que farão parte do portal corporativo.

Segundo o DELPHI GROUP (2002), uma empresa de consultoria especializada em desenvolvimento de taxonomias organizacionais, existem quatro etapas fundamentais no processo de desenvolvimento de uma taxonomia:

- 1) Desenvolvimento da estrutura da taxonomia.
- 2) Categorização dos assuntos do documento em uma estrutura hierárquica.
- 3) Apresentação da informação de forma que auxilie o usuário a localizar o que necessita.
- 4) Monitoramento de inclusões de novas informações, mantendo a estrutura do conhecimento.

A Figura 6 mostra essas etapas principais e as secundárias que subsidiam o desenvolvimento de uma taxonomia em softwares pesquisados pela empresa Delphi (2002).

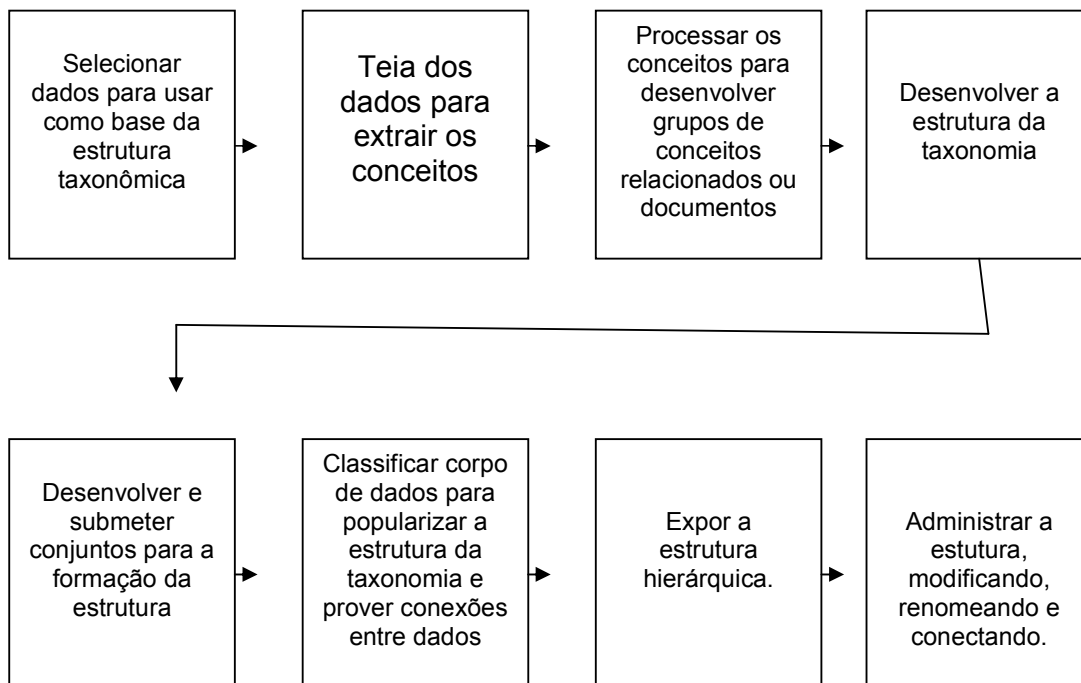


Figura 6: Processo geral de um software de taxonomia
Fonte: Delphi Group (2002)

De acordo com Holgate (2004), há quatro formas de se construir uma taxonomia:

- Iniciar de uma taxonomia pré-definida.
- Construir manualmente uma taxonomia.

- Construir automaticamente uma taxonomia.
- Uma combinação de automática e manual (híbrida).

Holgate (2004) afirma que a forma mais adequada depende de alguns critérios que variam em cada organização, como:

- O problema que a taxonomia está respondendo.
- O tipo e o alcance da informação corporativa.
- O volume do conteúdo.
- A disponibilidade dos especialistas da área para desenvolverem a taxonomia.

Woods (2004) explica que, em ambientes organizacionais, as taxonomias precisam:

- Fazer parte de um processo de gestão do conhecimento mais amplo.
- Estar relacionadas com arquiteturas de administração da informação, como portais, *datawarehousing*, etc.
- Estar relacionadas a um ambiente informacional integrado, procurando entender o fluxo informacional e semântico da organização.

Na construção de taxonomias corporativas, alguns critérios devem ser observados:

- Comunicabilidade: termos utilizados devem transparecer os conceitos carregados de acordo com a linguagem utilizada pelos usuários do sistema. Ex.: Cloreto de sódio (utilizado para especialistas) e sal (utilizado para leigos).
- Utilidade: apresentar somente os termos necessários. Ex. Frutas, sem especificar cada uma como maçã, pêra²¹.
- Estimulação: uso de termos que induzem o usuário a continuar a navegação pelo sistema.
- Compatibilidade: conter somente estruturas de campo que se está ordenando e que façam parte das atividades ou funções da organização.

(TERRA et al. 2005, p.3)

Yamaoka (2005) propôs uma metodologia para a construção de uma taxonomia corporativa, dentro de uma proposta de organização e representação do conhecimento em um portal, conforme Figura 7.

²¹ Nesse exemplo, o coeficiente de precisão será baixo, porque a linguagem de indexação empregada não foi suficientemente específica. (Nota do autor)

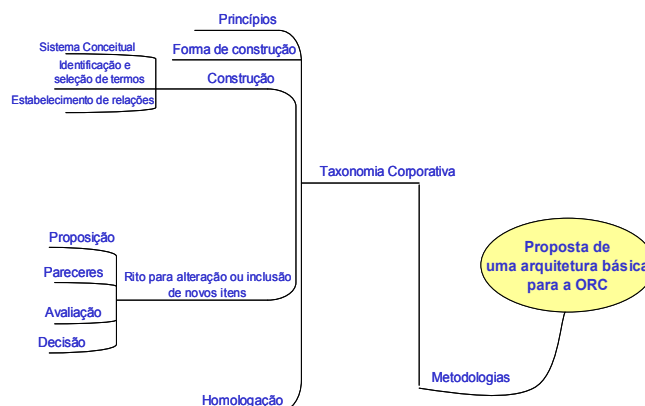


Figura 7: Arquitetura – Metodologias
Fonte: Yamaoka (2005)

A metodologia é dividida em cinco etapas principais:

- 1) Princípios
 - Visar à utilização como instrumento de apoio à utilização do conhecimento na instituição, através da navegação.
 - Respeitar a cultura da instituição, dando preferência aos termos usuais dos empregados.
 - Ser independente da estrutura organizacional da instituição.
 - Ter como base a necessidade de informação dos empregados.
 - Contemplar o comportamento dos usuários na busca da informação.
 - Refletir a necessidade de negócios e a necessidade informacional da instituição.
- 2) Forma de construção – tendo em vista os princípios para a construção da taxonomia, o processo de construção requer a participação das diferentes áreas da instituição, de forma a garantir a participação dos representantes e especialistas de cada área;
- 3) Construção – para a etapa de construção é necessária a concepção do sistema conceitual da taxonomia a ser concebida. Os princípios da metodologia para a construção da taxonomia, que estabelecem as necessidades que devem ser atendidas, podem direcionar a definição desse sistema conceitual.
- 4) Rito para alteração ou inclusão de novos termos – para garantir a manutenção contínua e manter sincronizadas as alterações com os interesses da instituição, é proposto para a arquitetura um rito que deve ser seguido para a alteração ou inclusão de novos termos na taxonomia corporativa. Os passos são: proposição, pareceres, avaliação e decisão.
- 5) Homologação – nesta etapa, busca-se estabelecer um processo de verificação da adequação da taxonomia em relação aos usuários e ao conteúdo que representa (YAMAOKA, 2005, p.134).

A formação de taxonomias beneficia-se da marcação dos documentos. A marcação é um passo importante, não só para garantir que mecanismos de busca encontrarão os documentos mais facilmente e os organizarão em classes, mas também para a disseminação dos documentos com base em regras de personalização. O conceito de metadados, ou dados que descrevem dados, base

das linguagens de marcação como HTML e XML, é central para a evolução da Internet, tanto em sua parte pública quanto na corporativa, representada pelos portais.

Os metadados provêm os contextos de significados necessários à informação e tornam os documentos mais facilmente localizáveis por mecanismos de busca, auxiliando-os a resolver diversas ambigüidades das línguas naturais que trazem ruídos ao processo de busca. A W3C²² define metadados como sendo “[...] informações localizadas na *Web*, inteligíveis por um computador” e, segundo Terra e Gordon (2002), “facilitando assim a localização dos documentos pelos mecanismos de busca independentemente de alguma informação prévia da localização física”.

Quando tratamos de ambientes *Web*, a definição de localização física não cabe, pois o local onde o documento está armazenado não interfere na sua localização. O que, de fato, precisa estar bem estruturado são as linguagens de marcação, os ‘caminhos’ que conduzem até a informação, como, por exemplo, autoria, assunto, área do conhecimento, entre outros.

Após o desenvolvimento de uma taxonomia, qualquer pessoa autorizada pode alocar as informações nos campos pré-determinados, tais como, *newsletter*, fale conosco, notícias, casos de sucesso, entre outros, fazendo com que a organização e a coerência no ambiente informacional sejam mantidas. A taxonomia, como qualquer outra forma de representação de informação e conhecimento, precisa acompanhar a natureza dinâmica e as variações lingüísticas que ocorrem no fluxo informacional das organizações e, conseqüentemente, estar em constante manutenção.

2.4 RELAÇÕES SEMÂNTICAS ADOTADAS EM MODELOS DE REPRESENTAÇÃO DO CONHECIMENTO

A linguagem natural associa um signo lingüístico a um significante (imagem acústica) e ao seu significado (conceito, imagem mental), representados no chamado triângulo de Ullmann (1959 *apud* BALDINGER, 1970) (Figura 7).

²² WorldWide Web Consortium

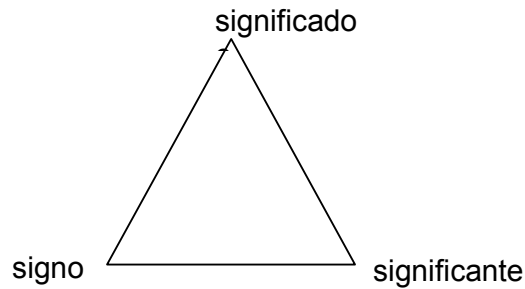


Figura 8: Triângulo de Ullmann
Fonte: Baldinger (1970)

Os sons da fala, imagens gráficas ou acústicas caracterizam o significante. O significado é uma idéia, conceito ou imagem mental.

No processo de estabelecimento da significação dos termos, significantes podem ter significados distintos (polissemia), assim como um significado pode ser representado por diferentes significantes (sinonímia).

A polissemia, segundo o Dicionário de Linguística, “é a propriedade que um signo linguístico tem de possuir vários sentidos”. O significante é portador de várias representações mentais. A homonímia acontece em termos que apresentam identidade de sons ou de forma, mas diversidade de significado. Porém, como aponta Baldinger (1970), uma mesma palavra com significações muito distintas, principalmente em áreas de especialidades, pode ser considerada como palavras diferentes. Essa constatação nos parece relevante, devido ao fato de uma taxonomia corporativa tratar informações advindas de diferentes setores, com processos distintos e que compartilham o portal como ambiente informacional.

Tanto a polissemia quanto a homonímia são geradoras de ambigüidade, entendida por Cintra, et al. (2002) “[...] como a possibilidade de uma comunicação lingüística prestar-se a mais de uma interpretação.” Como a linguagem natural se desenvolve nas interações entre os sujeitos, sem controle terminológico, os casos de sinonímia e polissemia são freqüentes. Mas então, como poderíamos entender a que significado o sujeito está se referindo em seu discurso? Segundo Baldinger (1970), é o contexto que determina o significado, a palavra se insere em uma frase, que apresenta uma contextualização mais ampla.

A sinonímia acontece quando dois termos designam um mesmo conceito. Cintra, et al. (2002) chamam de termos equivalentes entre, ao menos, duas

palavras; contudo, em alguns casos, essa relação de equivalência não se apresenta de forma absoluta. Lyons (1987, p.143) afirma que só podemos considerar duas palavras “[...] absolutamente sinônimas se, e somente se, tiverem a mesma distribuição e forem completamente sinônimas em todos os seus significados e contextos de ocorrência.” Cintra, et al. (2002) apontam a questão da sinonímia absoluta como um ponto de divergência entre estudiosos da língua; alguns negam a possibilidade de sua existência, enquanto outros usam a denominação ‘quase-sinonímia’. Poderiam ser citados como exemplo de quase-sinonímia os nomes populares e científicos que designam um mesmo objeto - gaivota (nome popular) e *Larus* (nome científico), termos usados em diferentes contextos e linguagens. A corrente que defende a impossibilidade da sinonímia leva em consideração que, em linguagens naturais, dois termos com a mesma significação dificilmente co-existem por muito tempo; sendo assim, sinônimos absolutos são raros.

A necessidade de controlar essas manifestações lingüísticas é fundamental no desenvolvimento de qualquer modelo de representação do conhecimento e organização da informação a fim de qualificar o processo de recuperação e representação.

Para entendermos as relações, é imprescindível a distinção entre termo e conceito. No âmbito da representação da informação, termo “é a palavra (ou expressão) utilizada para a inclusão temática de um item no sistema de informação e para recuperação posterior do mesmo assunto” (CAVALCANTI, 1978). Campos (2001), partindo da Teoria da Terminologia, apresenta termo como “[...] a unidade de comunicação que representa o conceito e pode ser constituído de uma ou mais palavras, uma letra, um símbolo gráfico, uma abreviação, uma notação.” Desse modo, para entendermos termo, é necessário conhecer a definição de conceito. Dahlberg (1978a), conforme citado, define conceito como “a unidade de conhecimento que surge pela síntese dos predicados necessários relacionados com determinado objeto e que, por meio de sinais lingüísticos, pode ser comunicado”. Seria, assim, a junção das características dos objetos ou fenômenos e, segundo Campos (2001), “um elemento de significação do termo”, representando um objeto na realidade empírica.

Sendo a linguagem natural suscetível no seu desenvolvimento e uso a ocorrências como polissemia e sinonímia, as linguagens documentárias surgem

como uma forma de uniformizar um vocabulário específico a fim de tornar possível o intercâmbio e a recuperação da informação.

Como visto, um termo representa um conceito, e conceitos são formados por características. Para Campos (2002), “a característica de um conceito pode ser também conceito, estabelecendo relações entre os conceitos que tenham uma característica em comum.” Assim, os termos são definidos uns em relação aos outros (CAMPOS, 2002). Definição compartilhada por Dahlberg (1978), que aponta a apresentação de determinadas características pelos conceitos como indutores à verificação das relações entre eles.

As relações semânticas são fundamentais no estabelecimento da significação em linguagens documentárias. Como mencionado, um termo pode apresentar diversos significados, o que se denomina polissemia, além de apresentar termos homônimos e equivalentes; todas essas manifestações precisam ser controladas nas linguagens construídas, para a recuperação ser eficaz.

A teoria da Classificação, aplicada nos sistemas tradicionais como Classificação Decimal Dewey e a Classificação Decimal Universal, utiliza as relações de gênero e espécie para a ordenação dos termos. Segundo Piedade (1977, p.20), gênero seria “[...] um conjunto de ‘coisas’ ou idéias, que possuem um ou vários atributos, predicados e qualidades em comum, podendo ser divididos em duas ou mais espécies.” As espécies “são os vários grupos resultantes da divisão de um gênero por determinada característica.” Com a identificação das características essenciais dos conceitos, essas relações são organizadas de forma hierárquica. Na organização hierárquica, de acordo com Araújo (2006, p.126), apenas um princípio pode ser aplicado por vez. O autor cita como exemplo um conjunto de cadeiras em que

usando a categoria quantidade, seriam separadas as grandes, das médias e das pequenas. Usando a categoria qualidade, pensada aqui como a cor, separar-se-iam as brancas, das azuis e verdes [...] Num sistema hierárquico, contudo, não podem ser utilizados esses [dois] princípios ao mesmo tempo, mas apenas um. Se for escolhida a categoria quantidade, então se terá, para o gênero cadeira, espécies grandes, médias e pequenas. Em cada espécie haverá cadeiras de diferentes cores. Se for a intenção, agora, separar por cores, terá de ser aplicado um novo princípio de classificação [...].

Pensando nessas limitações, Ranganathan desenvolveu a Classificação Facetada, cujo princípio é a construção de facetas. Facetas são definidas por Piedade (1977, p.22) como “a totalidade das subdivisões resultantes da aplicação de

uma única característica.” A construção da classificação por facetas permite uma análise ampla de assuntos, tornando possível a “simultaneidade dos critérios classificatórios” (ARAÚJO, 2006). Conforme mencionado em 2.3.1, as categorias fundamentais criadas por Ranganathan são: P (personalidade), M (matéria), E (energia), S (espaço-Local) e T (tempo); todas as facetas possíveis de um assunto seriam, segundo Ranganathan, manifestações de uma das categorias fundamentais.

No que se refere às relações conceituais, identificamos três classificações, a da Teoria Geral da Terminologia, a da Teoria do Conceito de Dahlberg (1978b) e a classificação de Cintra; et al. (2002). Não é possível afirmar que são classificações distintas, apresentam pontos de convergência, conforme veremos a seguir.

A Teoria Geral da Terminologia divide as relações conceituais em dois tipos:

- Relações Ontológicas: Acontecem entre o conceito e a realidade. Wüester (1981 *apud* CAMPOS, 2002, p.74) diz que “as relações ontológicas nascem do fato de elevarem-se a um nível de abstração as relações que existem na realidade entre os indivíduos (relações ônticas) fazendo-se destas relações individuais [...] conceitos de relação (acima, abaixo).” Campos (2002) cita um exemplo dado por Wüester (1981), segundo o qual, em uma carta geográfica, um pedaço de terra não poder ser, ao mesmo tempo, uma província e um cantão dessa província. Campos (2002, p.77) afirma que “as relações ontológicas são relações indiretas entre conceitos, porque resultam das propriedades que possuem os representantes dos conceitos.” Neste exemplo citado, na definição de província e cantão, um determinado pedaço de terra só poderia fazer parte de um desses conceitos.

As relações ontológicas são divididas em dois tipos: as relações de contato e as relações de causalidade.

Relações de contato – Considerada a mais importante relação ontológica, pode ser de duas espécies: relação de coordenação - relação parte-todo - e relação de causalidade. A relação de coordenação pode ocorrer entre dois ou mais conceitos. Quando acontece entre dois conceitos, no processo de comparação, existem quatro formas:

A subordinação partitiva ocorre quando um objeto individual é uma parte (subordinação partitiva) de outro objeto que é o todo (superordenação partitiva). [...] A relação de coordenação partitiva existe entre dois objetos que representam parte de um todo comum. [...] A relação interseção partitiva ocorre entre dois conceitos que, comparados, possuem alguma parte em comum. [...] A relação diagonal partitiva é aquela que se apresenta quando duas partes de um todo comum estão em relação, e esta relação não é de subordinação nem de coordenação. (CAMPOS, 2002, p.79)

A relação de causalidade deriva de “um elo sucessivo de causas” (WÜESTER, 1981 *apud* CAMPOS, 2002, p.79) e pode ser: de parentesco (genealógica [pai-filho]; ontogenética [ovo-larva]; de substância [urânio-rádio]. Também a relação de causalidade pode apresentar relações: material-produto (madeira → mesa); instrumental (instrumento - seu uso).

- Relações Lógicas: Também chamadas de relações de semelhança, de similaridade, de abstração e genéricas. Relações em que “um único e mesmo indivíduo pode representar cada um dos conceitos sucessivos de uma mesma cadeia de abstração”. Um exemplo é o conceito de cão, que pode ser ao mesmo tempo um São Bernardo, um mamífero e um ser vivo (WÜESTER, 1981 *apud* CAMPOS, 2002, p.74). Podem ser divididas em dois grupos:

Relações lógicas de comparação – Ocorrem entre dois conceitos e podem ser de quatro tipos, de acordo com Campos (2002), baseada em Wüester (1981 *apud* CAMPOS, 2002, p.76): **relação de hiponímia lógica** - “quando um conceito tem todas as características do outro conceito, e este último possui uma característica adicional; pode-se dizer que este conceito é uma espécie do outro, o gênero”; **relação de coordenação lógica** - “[...] quando dois conceitos analisados são específicos do mesmo termo genérico, distinguindo-se entre si por uma única característica.”; **relação de interseção lógica** - acontece “[...] quando são comparados dois conceitos, cuja intensão só é idêntica parcialmente, isto é, nem todas as características são as mesmas, como por exemplo, ensino-instrução”; por fim, **relação diagonal lógica** - “quando duas espécies de conceitos de mesmo gênero não estão ligados por relação de subordinação nem de coordenação, como, por exemplo, balão/foguete, navio/avião”.

Relações de combinação lógica – É assim definida por Wüester (1981 *apud* CAMPOS, 2002, p.76): “Consideremos três conceitos que não são ligados entre si, nem por uma relação de hiponímia lógica, nem por relação de coordenação lógica. Dois dentre eles podem então ser conceitos de partida que, por sua ligação (chamada também de combinação), constituem um terceiro conceito.”

As relações de combinação lógica são de três espécies: determinação, conjunção e disjunção. A determinação acontece quando um conceito aparece como suplementar a outro, auxiliando na sua compreensão e constituindo, desse modo, um terceiro conceito. O exemplo dado por Wüster (1981 *apud* CAMPOS, 2002, p.77)

é, *Import* = importar e *Kaufmann* = homem de negócios que resulta no conceito *Import-Kaufmann* = importador. A conjunção acontece quando dois conceitos se juntam para formar um terceiro; este, comum aos dois conceitos de partida. O exemplo é *Ingeenieur* = engenheiro e *Kaufmann* = homem de negócios que resulta, em alemão, em *Ingenieur-Kaufmann* = engenheiro-homem de negócios. Um *Ingenieur-Kaufmann* é tanto um engenheiro quanto um homem de negócios. A disjunção de conceitos, segundo Wüster (1981 *apud* CAMPOS, 2002, p.77), acontece quando “as extensões dos dois conceitos, ou seja, seus específicos são reunidos. Um exemplo: indo-europeu. Daí resulta o genérico comum dos conceitos de partida”.

A Teoria do Conceito de Dahlberg (1978b) apresenta dois tipos de relações conceituais: as relações formais, que se baseiam na comparação dos conceitos e as relações materiais, que tratam do conteúdo das características dos conceitos. De acordo com Dahlberg (1978b), as relações materiais podem ser dos seguintes tipos:

- **Relações Abstrativas ou Genéricas (gênero-espécie):** Acontecem quando dois conceitos possuem as mesmas características, e um em relação ao outro possui uma característica adicional, surgindo uma hierarquia. Dahlberg (1978b, p.15) cita como exemplo:

jogo

jogo de tabuleiro

jogo de tabuleiro para dois jogadores

jogo de xadrez

- **Relações Partitivas:** Relação existente entre “[...] o conceito de um todo e qualquer uma das partes desse todo”; as partes podem ainda ser subdividas em outras partes ou sub-elementos. (DAHLBERG, 1978b).

- **Relações de Oposição:** Podem acontecer de duas maneiras: contraditoriedade e contrariedade. Segundo Dahlberg (1978b), a contrariedade seria negação do conceito, branco/não branco; já a contraditoriedade seria uma terceira possibilidade, branco/verde. Relação que geralmente acontece entre as propriedades dos objetos.

- **Relações Funcionais (também chamadas sintagmáticas):** São caracterizadas como relações resultantes de um processo. Os conceitos podem ser reduzidos às etapas do processo, sendo uma relação importante para o momento de definição dos conceitos e no processo de análise-síntese dos conteúdos temáticos.

Cintra et al. (2002) definem as relações semânticas como sendo de dois tipos: hierárquicas e não-hierárquicas ou seqüenciais.

Para o entendimento das relações hierárquicas, as autoras estabelecem a definição de ordem e subordinação, que determinam esse tipo de relação.

A ordem deve ser observada como uma superordenação que consiste na possibilidade de subdivisão de uma noção hierárquica mais alta em certo número de noções de nível inferior, chamadas noções subordinadas. É este processo de subdivisão que se denomina subordinação. Inversamente, a noção subordinada é a noção que, num 'sistema hierárquico', pode ser agrupada com uma ou mais noções do mesmo nível (noções coordenadas entre si), para formar uma noção de nível superior, ou seja, uma superordenação. (CINTRA et al. 2002, p.51)

É nesse processo que as relações hierárquicas se estabelecem. Um conceito superordenado divide-se em conceitos subordinados, com relações coordenadas entre eles, conforme Figura 8. É o equivalente às relações gênero-espécie.

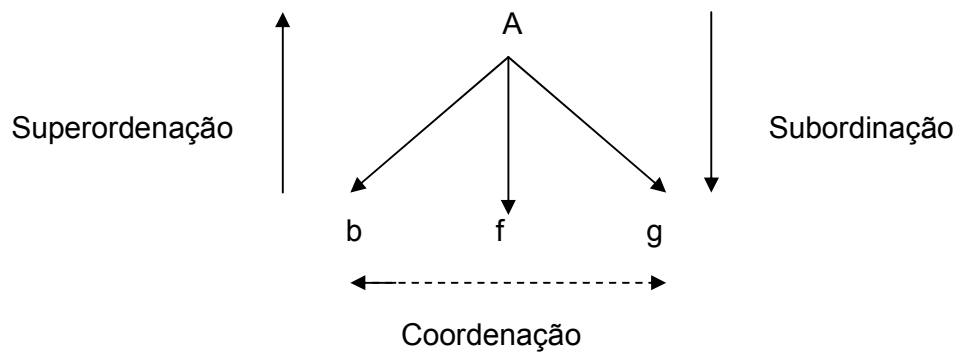


Figura 9: Esquema de relações hierárquicas
Fonte: Cintra et al. (2002, p.51)

Segundo Cintra et al. (2002), as relações hierárquicas podem ser de três tipos²³:

- Relações genéricas: a superordenação parte das diferenças para as semelhanças; a subordinação parte das semelhanças para as diferenças. Por exemplo, o conceito 'meios de transporte' subdivide-se nos tipos, automóvel, motocicleta, trem, etc. Em relação aos seus diferentes tipos, 'meios de transporte' é a noção superordenada; desse modo, podemos dizer que trem é um tipo de meio de transporte e que trem e motocicleta são noções coordenadas.

²³ No nosso entender as relações genéricas e específicas fazem parte de um único tipo de relacionamento hierárquico.

- Relações específicas: são relações hierárquicas subordinadas que compartilham as mesmas características da noção (conceito) superordenada, mas apresentam uma característica a mais que as diferenciam.
- Relações partitivas: acontecem quando a noção superordenada refere-se a um objeto considerado como um todo e as noções subordinadas partes dele. Por exemplo, em relação a trem, os vagões são noções específicas partitivas, sendo o trem a noção do todo e os vagões uma noção de parte, chamada de relação coordenada na Teoria Geral da Terminologia.

As relações não-hierárquicas ou seqüenciais são, de acordo com Cintra et al. (2002), as que não se submetem a uma hierarquia, apresentando entre si contigüidade espacial ou temporal. A autora fala sobre a dificuldade de definir essas relações, pois usualmente são denominadas associativas, e qualquer natureza de relação seria associativa. Dessa forma, elas seriam definidas pelo universo referencial de cada indivíduo; porém, na construção de linguagens documentárias, é necessário ter controle sobre esse referencial. “Tais referências só podem estar assentadas em princípios funcionais, reconhece Dahlberg, para quem um relacionamento funcional é ‘aquele em que um termo que denote atividade ou operação se liga, conceitualmente, a uma entidade ou propriedade’.” (IBICT, 1984, p.31 *apud* CINTRA et al., 2002, p.63). Para isso, é necessário mapear e definir os conceitos de uma área específica, objetivando atingir a consistência nos procedimentos. Cintra et al. (2002) apresentam algumas possibilidades de combinação entre conceitos de natureza não-hierárquica ou seqüenciais, conforme ilustrado no Quadro 5.

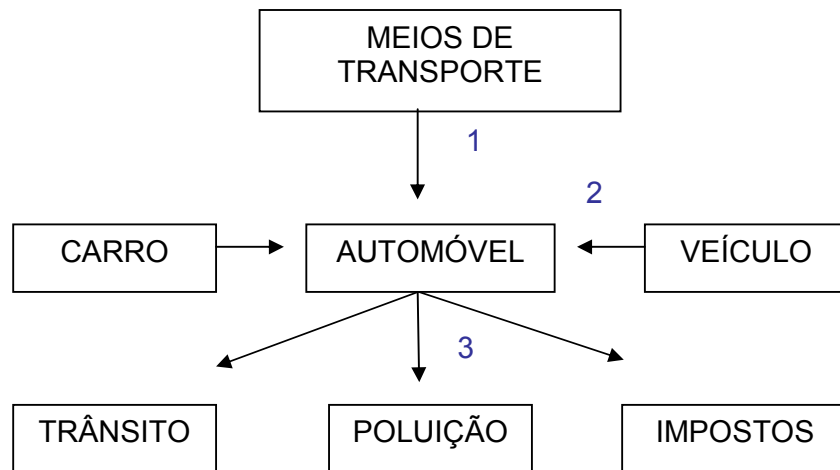
RELAÇÃO	EXEMPLO	
Relação de Atribuição	Economia	Nível de atividade econômica
Disciplina ou campo de estudo ou fenômenos estudados	Entomologia	Insetos
	Estética	Beleza
	Pacifismo	Paz
Processo ou operação/seu agente ou instrumento	Controle da temperatura	Termostatos
	Automação	Computadores
	Iluminação	Lâmpadas
Relação de influência	Política Monetária	Inflação
Matéria-prima/produto	Bauxita	Alumínio
Ação/resultante da ação	Tecelagem	Tecidos
	Pintura (arte)	Murais
Causalidade ou causa/conseqüência	Crescimento econômico	Desenvolvimento econômico
Efeito/causa	Febre	Infecção
Atividade/agente	Tabagismo	Fumo
Opostos	Vida	Morte
Ação/seu paciente	Extradição	Criminosos
	Pesca	Pescado
Atividade/produto	Tear	Tecido
Coisa ou atividade/suas propriedades ou agentes	Venenos	Toxidade
	Criança Superdotada	Inteligência
Atividades complementares	Compra	Venda
Associação implícita	Balanço de pagamento	Comércio internacional

Quadro 5: Tipologias de relações seqüenciais

Fonte: Cintra et al. (2002, p.64)

É possível verificar, pelo Quadro 5, a diversidade de relacionamentos possíveis resultantes das relações seqüenciais.

Considerando que as relações semânticas hierárquicas, de equivalência e seqüenciais abrangem todas as possibilidades de relacionamentos entre termos, elas são as bases para a análise na construção de taxonomias. Essas relações podem ser exemplificadas da seguinte forma:



1 Relação Hierárquica

2 Relação de Equivalência

3 Relação Sequencial de Associação Implícita

Figura 10: Relações seqüenciais

Fonte: Do autor

Podemos constatar, nas três classificações, várias similaridades. Nas relações conceituais estabelecidas na Teoria Geral da Terminologia podemos perceber, tanto nas relações ontológicas quanto nas lógicas, a caracterização das relações hierárquicas (relação de contato), de equivalência (relações lógicas de comparação) e as seqüenciais (relações de combinação lógica). Da mesma forma, na Teoria do Conceito, as relações genéricas e partitivas são tipos, como afirma Cintra; et al. (2002) de hierarquias; e, as relações funcionais caracterizam-se como seqüenciais. Nas três formas de classificação, a que nos parece mais adequada à sistematização das relações conceituais em uma taxonomia corporativa é a estabelecida por Cintra; et al. (2002). As relações hierárquicas e, principalmente as não-hierárquicas ou seqüenciais possibilitam a representação da diversidade de relacionamento do ambiente corporativo, possibilitando o desenvolvimento de uma lógica própria a cada ambiente, um dos objetivos da taxonomia.

3 PROCEDIMENTOS METODOLÓGICOS

Nos próximos tópicos, apresentaremos a caracterização da pesquisa e os procedimentos metodológicos utilizados no processo de coleta e análise dos dados. Procedimentos estes que objetivam fornecer coerência e validade científica à pesquisa.

3.1 TIPO DE PESQUISA

Como o campo estudado não dispõe de um conjunto sistematizado de conhecimentos consolidados, a forma mais adequada de proceder é a da exploração e análise dos conceitos existentes na literatura, com o objetivo de realizar inferências, caracterizando assim o cunho qualitativo e exploratório deste estudo.

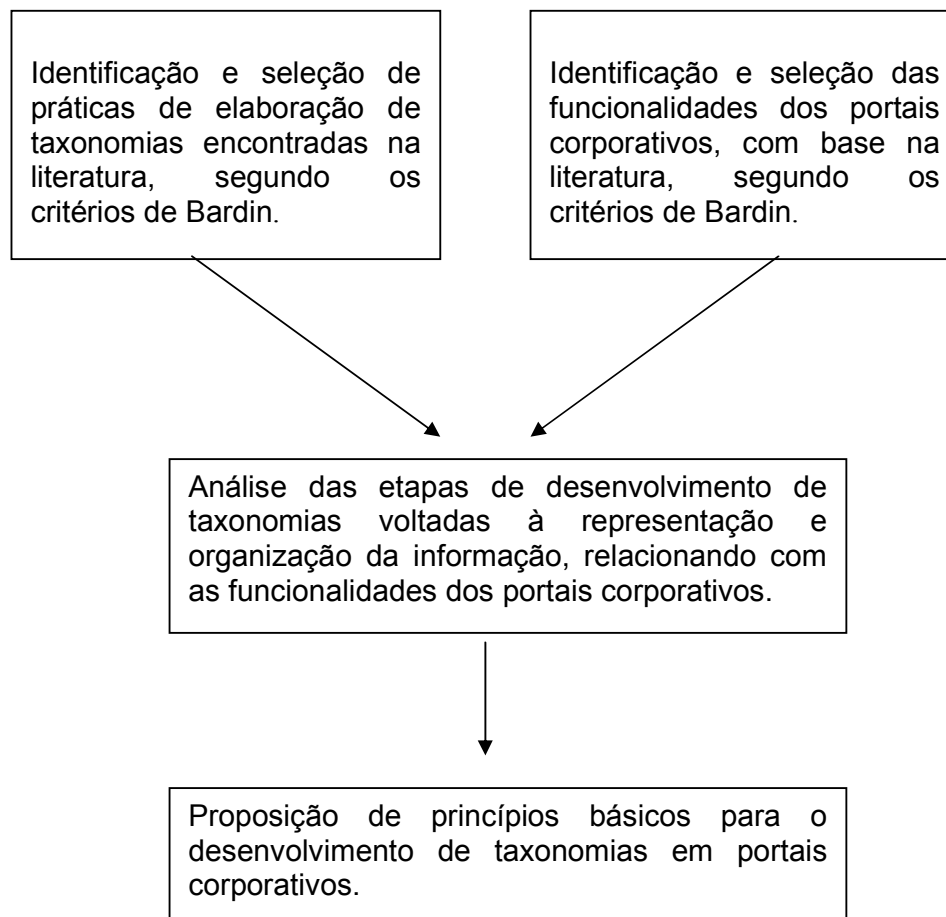
Esta pesquisa é, portanto, de natureza qualitativa, pois realiza uma análise da literatura e propõe diretrizes para o desenvolvimento de uma taxonomia, sem a pretensão de efetuar medições ou fazer uso de técnicas estatísticas. De acordo com Deslantes et al. (1994, p.21), a pesquisa qualitativa “[...] trabalha com o universo de significados, motivos, aspirações, crenças, valores e atitudes, o que corresponde a um espaço mais profundo das relações, dos processos e dos fenômenos que não podem ser reduzidos à operacionalização das variáveis”. Caracteriza-se igualmente como exploratória, pois tem como foco proporcionar um conhecimento mais amplo do campo estudado. Um estudo exploratório, segundo Gil (2002), “[...] tem como objetivo principal o aprimoramento de idéias ou a descoberta de intuições” e acontece, principalmente, em áreas do conhecimento em que há pouco material sistematizado, tendo assim, um planejamento mais flexível, porém não descompromissado.

3.2 ETAPAS DA PESQUISA

A pesquisa apresenta quatro momentos distintos. O primeiro diz respeito à identificação e seleção de práticas de elaboração referentes ao desenvolvimento de taxonomias. O segundo deles refere-se à identificação e seleção de funcionalidades discutidas na literatura e relacionadas à representação e organização da informação em portais corporativos. O terceiro momento apresenta a análise das etapas de

desenvolvimento de taxonomias, adotando como critério de análise as funcionalidades definidas no segundo momento. Finalmente, o quarto trata do estabelecimento de princípios para a construção da taxonomia em portais corporativos.

Esses momentos são sintetizados no esquema a seguir, sendo cada um deles detalhadamente explicados nas seções posteriores.



3.2.1 IDENTIFICAÇÃO E SELEÇÃO DAS PRÁTICAS DE ELABORAÇÃO DE TAXONOMIA

As práticas de elaboração de taxonomia constituem-se no *corpus* desta pesquisa, cuja natureza é bibliográfica. Sua coleta apóia-se na técnica de análise de conteúdo definido por Bardin (2004) para a fase de pré-análise. Segundo a autora, a pré-análise “[...] tem por objetivo tornar operacionais e sistematizar as idéias iniciais,

de maneira a conduzir a um esquema preciso do desenvolvimento das operações sucessivas, num plano de análise”. (BARDIN, 2004, p.89). Entre os critérios definidos por Bardin (2004) para esta fase, utilizamos:

- a) Regra da exaustividade - O material coletado é constituído de artigos de periódicos da área de Ciência da Informação, com texto integral indexados no Portal Capes, entre os anos de 2000 e 2005, que apresentam práticas de elaboração de taxonomias. A escolha do Portal Capes se deve ao fato do mesmo disponibilizar produção científica relevante, em termos quantitativos e qualitativos, da área de Ciência da Informação. Além de artigos, outros materiais bibliográficos, com a mesma metodologia, também foram utilizados. A escolha somente de documentos dos últimos 5 anos se deve ao fato de a aplicação de taxonomias para a recuperação da informação em portais corporativos ser um tema recente na literatura.
- b) Regra da pertinência – A partir da recuperação dos artigos, foram descartados os que não se enquadraram na regra de pertinência, segundo a qual “os documentos retidos devem ser adequados, enquanto fonte de informação, de modo a corresponderem ao objetivo que suscita a análise”. (BARDIN, 2004, p.91). Artigos que não definiam as etapas e os procedimentos metodológicos para o desenvolvimento de taxonomias foram descartados.

Além dos critérios de exaustividade e pertinência, o *corpus* da pesquisa teve como delimitador o idioma. Somente foram recuperados artigos nos idiomas, inglês, português e espanhol. As palavras-chave para a busca foram limitadas a três: ‘metodologia e taxonomia’, ‘taxonomia corporativa’ e ‘taxonomia’.

3.2.1.2 PRÁTICAS DE ELABORAÇÃO DE TAXONOMIAS

Com base na pesquisa realizada na literatura, seis práticas de elaboração de taxonomias foram selecionadas.

A primeira prática de elaboração de taxonomia (Prática de Elaboração 1) foi criada pela empresa *Earley & Associates*, especializada em consultorias na área de organização e disponibilização da informação e do conhecimento. No artigo intitulado 'Desenvolvendo taxonomias para empresas', de 2003, foi possível identificar a prática de elaboração de taxonomia proposta por esta empresa. Essa prática tem com premissa o fato de toda taxonomia precisar ter um significado para a equipe que fará uso dela. Para tanto, é sugerida a formação de grupos de trabalho para a construção da taxonomia, compostos por taxonomistas ou especialistas.

A segunda prática de elaboração de taxonomia (Prática de Elaboração 2) foi desenvolvida pela National Aeronautics and Space Administration (NASA), americana, e teve origem em um ato governamental, indicando o desenvolvimento de padrões e diretrizes para a categorização da informação em formato eletrônico. A taxonomia prevista por esta prática de elaboração tem como objetivo integrar os diferentes centros de especialidades que, por utilizarem linguagens distintas, encontravam dificuldades na comunicação. Dutra e Busch (2003) apresentam prática de elaboração dividindo-a em duas fases. A primeira refere-se à construção da taxonomia *alpha*, em que é permitida a participação de todos os usuários. Para seu desenvolvimento, são definidas diretrizes mais gerais no que diz respeito aos termos, e o produto final é uma taxonomia mais ampla. A segunda fase tem por objetivo elaborar a chamada taxonomia *beta*, desenvolvida por pequenos grupos de especialistas em diferentes assuntos. O resultado é uma taxonomia com níveis mais profundos de relações entre os termos.

A terceira prática de elaboração de taxonomia (Prática de Elaboração 3) encontrada é de autoria de Kremer; Kolbe e Brenner (2005). Estes autores discorrem sobre a importância do que denominam 'gerenciamento da terminologia', apresentando o glossário e a taxonomia como duas ferramentas resultantes desse processo. Para eles, a taxonomia é um esquema de classificação que utiliza o glossário (explicitação das relações entre os termos) combinado a ela para proporcionar uma recuperação eficiente.

A quarta prática de elaboração de taxonomia (Prática de Elaboração 4) é definida por Yamaoka (2005), que desenvolveu uma proposta de arquitetura para a organização e representação do conhecimento (ORC). A proposta é pautada em oito elementos: objetivos, princípios, variáveis, normas, repositório, metodologias, base de dados terminológica e escritório de ORC. A prática de elaboração inclui a determinação de metadados e da taxonomia corporativa. Para a etapa de construção da taxonomia, é estabelecido um sistema conceitual - a identificação e seleção de termos e o estabelecimento das relações entre eles.

A quinta prática de elaboração de taxonomia (Prática de Elaboração 5) é a do Montague Institute (2001), que divide a taxonomia em duas partes: a estrutura e a aplicação. A estrutura consiste nas categorias ou termos e as relações entre eles; a aplicação, por outro lado são as ferramentas de navegação que auxiliam o usuário a localizar a informação, uma dependente da outra.

A sexta e última prática de elaboração de taxonomia (Prática de Elaboração 6) é descrita pela empresa de consultoria Kcurve. A autora, Márcia Morante, trabalha com desenvolvimento de taxonomias para organizações. A prática de elaboração foi localizada em uma apresentação de Morante em 2003, a qual especifica as etapas no processo de desenvolvimento de taxonomias.

3.2.2 IDENTIFICAÇÃO E SELEÇÃO DAS FUNCIONALIDADES DOS PORTAIS CORPORATIVOS

As funcionalidades referentes ao processo de representação e organização da informação em portais corporativos constituem-se nas categorias utilizadas para classificação e análise do corpus. Para defini-las, adotamos o processo de categorização, definido por Bardin (2004) como “[...] uma operação de classificação de elementos constitutivos de um conjunto, por diferenciação e, seguidamente, por reagrupamento segundo o gênero (analogia), com os critérios previamente definidos”. Para a realização da classificação, é necessário o estabelecimento de categorias, que reunirão os elementos comuns, seguindo algumas determinações de acordo com Bardin (2004, p. 114):

- a) Exclusão mútua – cada elemento só pode fazer parte de uma categoria, e essas devem ser construídas de tal forma a não permitir

que dois elementos tenham características possíveis de classificá-los em mais de uma categoria.

- b) Homogeneidade – a organização das categorias deve seguir um único princípio de organização. Diferentes aspectos da análise devem ser desdobrados em outras tantas categorias.
- c) Pertinência – as categorias devem ser condizentes ao quadro teórico que norteia a pesquisa, os objetivos e o problema proposto, assim como ao material a ser analisado.
- d) Objetividade e fidelidade – a subjetividade inerente a quem está codificando o material deve ser minimizada por categorias bem estabelecidas.

Essas determinações foram aplicadas na literatura revisada, resultando na identificação da categoria **categorização**, por nós definida de funcionalidade, que orientou a análise do material a fim de propor recomendações para o desenvolvimento de uma taxonomia em portais corporativos.

3.2.2.1 FUNCIONALIDADES DOS PORTAIS CORPORATIVOS

As diferentes funcionalidades são requisitos, considerados pelos autores, indispensáveis no ambiente dos portais corporativos. A aplicação das determinações, definidas por Bardin (2004) e descritas no item anterior, resultou na identificação da funcionalidade **categorização** como a variável de análise relacionada à representação e organização da informação em portais corporativos. Esta funcionalidade é descrita na literatura da área de portais corporativos por quatro autores, conforme apresentado no Quadro 6. O trabalho de cada um dos autores selecionados no estabelecimento da funcionalidade foi tratado no item 2.2.

Funcionalidades relacionadas à categorização da informação	Terra e Gordon (2002)	Eckerson (1999 <i>apud</i> DIAS, 2001)	Carvalho e Ferreira (2005)	Detlor (2000)
Categorização				

Quadro 6: Funcionalidades dos portais relacionadas à organização e representação da informação
Fonte: Do autor

A funcionalidade **categorização** é a variável central de análise para esse trabalho. A literatura da área de portais corporativos mostra que o conceito de categorização envolve tanto aspectos da construção de taxonomia (representação do conhecimento) como da aplicação da taxonomia na representação da informação e organização da informação recuperada.

Na Ciência da Informação, a categorização diz respeito ao estabelecimento de classes gerais, as chamadas facetas, cujo agrupamento forma as categorias fundamentais, conforme definido pela teoria da Classificação Facetada de Ranganathan. Nessas classes gerais, estão organizados os conceitos em renques²⁴ e cadeias²⁵.

O processo de categorização é descrito por todos os autores dispostos no Quadro 6 como uma funcionalidade fundamental, pois é por ela que a recuperação da informação é viabilizada. Com base no reconhecimento das informações existentes no ambiente do portal, são criadas as categorias que façam sentido para a organização. Nessas categorias, são inseridas as informações pertinentes a elas.

Terra e Gordon (2002) entendem a categorização como um processo paralelo à taxonomia; no entanto, autores como Martinez et al. (2004) e Adams (2000) vêem a taxonomia como o próprio processo de categorização. Segundo estes autores, na taxonomia são criadas estruturas gerais (categorias) com o objetivo de organizar a informação.

No intuito de completar a noção de categorização apontada por estes autores, buscamos subsídios na área da Ciência da Informação, fazendo o cruzamento dos conceitos quando necessário. Assim, Campos (2002); Martinez et al. (2004); Cintra et al.(2002); Gardin et al. (*apud* CINTRA et al., 2002) e Piedade (1977) identificam as seguintes etapas na construção de modelos de representação do conhecimento que podem, no nosso entender, ser aplicadas a taxonomias.

1. Estabelecimento das categorias gerais.
2. Coleta dos termos.
3. Análise dos termos selecionados.
4. Controle da diversidade de significação.
5. Construção dos relacionamentos semânticos.

²⁴ “Renques são classes formadas a partir de uma única característica de divisão, formando séries horizontais”. (CAMPOS, 2001)

²⁵ “Cadeias são séries verticais de conceitos em que cada conceito tem uma característica a mais ou a menos conforme a cadeia seja descendente ou ascendente”. (CAMPOS, 2001)

Na primeira etapa, o estabelecimento das categorias gerais na construção da taxonomia, assim como acontece no estabelecimento dos assuntos de um documento, é um processo que precisa levar em consideração o contexto e os objetivos a serem atingidos. Terra e Gordon (2002) apontam a necessidade da criação de categorias com o fim de facilitar a posterior representação das informações da organização, fazendo sentido aos seus usuários. Nessa etapa de definição de categorias, é necessário primeiramente identificar quais as grandes áreas do conhecimento que serão representadas. Em seguida, procede-se à coleta dos termos que serão incluídos nessas categorias. Na passagem da segunda para a terceira etapa, Gardin et al. (*apud* CINTRA et al., 2002) alertam que os elementos descritores precisam estar devidamente filtrados e depurados, ou seja, é preciso haver uma seleção dos termos. Quanto à quarta etapa, eles salientam sobre a necessidade do controle da significação dos termos escolhidos. Nesse sentido, Cintra et al.(2002) apontam a polissemia e a homonímia como geradores de ambigüidades necessitando assim serem controladas. E, por fim, em relação à construção dos relacionamentos semânticos, não basta somente selecionar, avaliar e controlar a significação dos termos. De acordo com Campos (2002, p.50), é importante “[...] estabelecer suas posições relativas, o que se obtém por meio da determinação das relações que as associam”.

No que se refere ao aspecto da aplicação da taxonomia no processo de representação da informação para recuperação da informação, Terra e Gordon (2002) apontam a indexação como processo interligado à categorização. Eckerson (1999 *apud* DIAS, 2001) segue esse mesmo raciocínio quando enquadra o processo de categorização na funcionalidade ‘Classificação’ e a aponta como a capacidade de o portal indexar e organizar as informações da empresa. Esse autor entende que a categorização engloba o processo de indexação, adicionando a este a organização dos termos nas relações mais pertinentes. Cavalcanti (1978) define indexação como o ato, processo ou resultado da elaboração de entradas para um índice. O processo de indexação engloba a análise e seleção dos termos de um documento e a tradução para uma linguagem controlada. Piedade (1977) relata que, no processo de indexação, a identificação dos termos ou conceitos é o primeiro passo para representar um documento.

Carvalho e Ferreira (2005) referem-se à categorização em portais, como um elemento de representação, “[...] que é uma função do subsistema de entrada do

SRI (Sistema de Recuperação da Informação) responsável pela tradução dos documentos para um sistema de classificação”. A categorização, por meio de uma metodologia, permite a seleção, análise, relacionamento e organização dos termos, apresentando assim o processo de indexação.

Detlor (2000) aponta a alocação da informação em categorias como a melhor forma de permitir a acessibilidade dos usuários a elas.

A categorização, referente ao aspecto da organização da informação recuperada é apresentada por Detlor (2000) como a forma de tornar mais simples e fácil esse processo para o usuário. Eckerson (1999 *apud* DIAS, 2001) também mostra a necessidade da criação de categorias condizentes com as necessidades da organização e de fácil compreensão, apresentando a informação disposta nessas categorias no momento da recuperação. Segundo o autor, isso facilitaria uma pesquisa intuitiva, na qual o usuário perceberia mais claramente os assuntos e suas relações naquele ambiente. Matinez et al. (2004) também indicam a taxonomia como responsável pela criação de estruturas gerais para a organização da informação. Esse ponto de vista é igualmente compartilhado por Terra e Gordon (2002) para quem esse processo coloca o usuário “[...] em contato com tópicos e categorias relacionadas não previstas anteriormente”. Dessa maneira, é possível verificar que esses autores interpretam a aplicação da taxonomia não somente na representação da informação como também na organização da informação recuperada.

Com base no exposto, as práticas de elaboração de taxonomia identificados no corpus são analisados sob a ótica da funcionalidade ‘categorização’ nos seguintes aspectos:

a) Construção da taxonomia (representação do conhecimento):

- estabelecimento das categorias gerais;
- coleta dos termos;
- análise dos termos selecionados;
- controle da diversidade de significação;
- construção dos relacionamentos semânticos.

b) Aplicação da taxonomia na:

- representação da informação (classificação ou indexação);
- organização da informação recuperada do acervo.

4 ANÁLISE DAS PRÁTICAS DE ELABORAÇÃO DE TAXONOMIAS

Neste capítulo, é apresentada a análise dos dados da pesquisa, objetivando a elaboração de proposições que subsidiem a construção de taxonomias corporativas.

Com base na pesquisa realizada na literatura, seis práticas de elaboração foram localizados. Essas práticas são analisadas, tomando por base a principal funcionalidade de uma taxonomia: a categorização. O conceito de categorização, exposto anteriormente, envolve os aspectos de construção da taxonomia e sua aplicação na representação da informação e na organização da informação recuperada.

A seguir, verificaremos como as práticas de elaboração abordam esses aspectos.

A) PRÁTICA DE ELABORAÇÃO 1

A primeira prática de elaboração (Prática de elaboração1), criada pela empresa *Earley & Associates*, tem como premissa o fato de toda taxonomia precisar ter um significado para a equipe que fará uso dela. Para tanto, é sugerida a formação de grupos de trabalho para a construção da taxonomia, compostos por taxonomistas ou especialistas.

Foram identificadas cinco etapas que compõem a prática de elaboração, a saber:

1º Verificar as necessidades informacionais dos usuários – Por meio do que os autores chamam de ‘autoria do conhecimento’, questiona-se sobre quais informações são indispensáveis para o desenvolvimento do trabalho do grupo.

2º Realizar testes de busca – Tomando por base *sites* considerados úteis e fáceis de navegar pelos usuários, identificam-se as preferências na forma de navegação e classificação dos termos.

3º Buscar termos de classificação – Baseando-se na análise da organização pessoal de cada usuário, promove-se uma discussão a respeito de quais as possíveis formas de classificar determinadas informações, procurando evidenciar os diferentes significados de um mesmo termo para o grupo.

4º Discutir termos da taxonomia – Com base nas discussões sobre as informações fundamentais da organização, levantam-se, junto ao grupo, termos que definam essas informações, apresentadas como um quebra-cabeça a ser montado, chegando a uma taxonomia representativa daquele grupo de usuários.

5º Testar a taxonomia – Realiza-se um teste-piloto com amostras da informação organizada anteriormente na estrutura da taxonomia.

Os autores desta prática ainda recomendam que a taxonomia deva ser revisada a fim de acompanhar mudanças e obsolescências de alguns termos.

Quanto à construção da taxonomia

Na etapa 1, o estabelecimento das categorias gerais é facilitado no trabalho cooperativo, chamado de 'autoria do conhecimento', evidenciando a necessidade de essas categorias terem um significado para o grupo. Para tanto, o conhecimento da área é fundamental. Esta etapa, portanto, vai ao encontro do que Terra e Gordon (2002) afirmam sobre a necessidade de a taxonomia fazer sentido a quem irá utilizá-la.

A etapa 2 complementa a identificação das categorias pela consulta às preferências nas formas de navegação e classificação em *sites* utilizados pelo grupo. Este procedimento também é salientado como importante por Terra e Gordon (2002) quando apontam a preocupação de averiguar as preferências dos usuários em termos de navegação.

Enquanto a etapa 2 parte da análise dos *sites*, a etapa 3 adota como parâmetro de definição de categorias a organização pessoal de cada integrante do grupo. Nesta etapa, ainda é ressaltada a necessidade de verificação do aparecimento de diferentes termos designativos de uma mesma categoria, isto é, aborda o tema da polissemia no nível categorial. No entanto, apenas reconhece a polissemia, sem especificar de que forma os diversos significados levantados pelo grupo serão escolhidos, descartados ou organizados na taxonomia. Além da variação no campo da significação, é importante também a identificação da variação terminológica para um mesmo conceito contemplado pela categoria, ou seja, a sinonímia, item não apontado na prática de elaboração.

A etapa 4 é dedicada à coleta dos termos e sua inclusão nas categorias da taxonomia. Apesar de sua importância, esta etapa é descrita de forma superficial. Não há evidência de que serão também consideradas aqui as questões de

polissemia e sinonímia. Além disso, a expressão 'montar um grande quebra-cabeça' não explicita como os termos e suas variações semânticas e lexicais serão organizados. Da forma como está exposta, a etapa 4 dificilmente encontrará um consenso entre os componentes do grupo. Nesta etapa, os autores da prática de elaboração afirmam também que, com base na discussão sobre os termos, chega-se a uma taxonomia representativa daquele grupo. No entanto, não há uma especificação de como esse processo ocorre. A determinação de critérios de cunho lingüístico provavelmente orientaria metodologicamente as decisões, uma vez que se trata de questões envolvendo o estabelecimento da estrutura conceitual da taxonomia.

Quanto à aplicação da taxonomia

A indicação, na prática de elaboração, da realização de testes de busca na etapa 5 mostra que a taxonomia terá a função de ser um instrumento de recuperação da informação. Nesse sentido, conforme salienta Dutra e Busch (2003), ela também deve ter como objetivo servir de referência para o entendimento dos assuntos e relações, visando à recuperação. Assim, por meio da taxonomia, o usuário poderá criar estratégias de busca e compreender ou conhecer as relações entre os diferentes termos representativos dos processos organizacionais. Conway e Sligar (2002) dizem, a este respeito, que a estruturação das informações precisa ficar visível ao usuário final, contribuindo para que, nesse processo de busca, novas informações sejam encontradas.

Em uma taxonomia, o teste-piloto é o momento em que a taxonomia é publicada. Com base nas representações construídas, organizam-se as informações no portal corporativo.

B) PRÁTICA DE ELABORAÇÃO 2

A prática de elaboração apresentada por Dutra e Busch (2003), desenvolvida para a NASA, orienta a construção da taxonomia em duas fases. A primeira fase - definição de alcance da taxonomia - pressupõe a participação do usuário na definição do escopo geral da taxonomia, objetivando atender às necessidades informacionais do grupo. A segunda fase - construção da taxonomia - parte desse

escopo mais amplo para as áreas de especialidades, objetivando aprofundar e especificar os termos e suas relações.

1ª fase - Definição de alcance da taxonomia:

- a) apresentar proposta de taxonomia;
- b) entrevistar usuários e *experts* no assunto;
- c) recolher exemplares da documentação produzida e de vocabulários controlados existentes;
- d) analisar informações coletadas.

2ª fase - Construção da taxonomia:

- a) desenvolver uma taxonomia *alpha*;
- b) apresentar os resultados à comunidade de interesse e proporcionar uma realimentação;
- c) revisar e alcançar um consenso na taxonomia *alpha*;
- d) desenvolver a taxonomia *beta*;
- e) apresentar a taxonomia à comunidade de interesse;
- f) revisar e alcançar um consenso na taxonomia *beta*.

Para efeito desta pesquisa, aprofundaremos apenas a fase de construção da taxonomia que pressupõe o desenvolvimento das taxonomias *alpha* e *beta*. A taxonomia *alpha* é a que determina a estrutura mais ampla da taxonomia, dois ou três níveis mais gerais, que podem ser definidos, de acordo com Dutra e Bush (2003), pelos seguintes mecanismos:

Inspiração: Ponto de vista de um consultor individual ou participante do projeto.

Indução: Análise da coleção de documentos utilizando ferramentas automáticas.

Dedução: Aplicação de um padrão existente, abordagem ou modelos.

Síntese: Combinação de abordagens.

O mecanismo de inspiração acontece quando consultores ou participantes do projeto, com base em suas observações a respeito das informações que circulam na organização, instituem as categorias consideradas representativas. A indução é realizada com base em ferramentas automatizadas que analisam a ocorrência dos termos nos documentos e concluem quais os mais significativos no contexto. A dedução parte das representações já existentes na organização, sejam padrões, abordagens referentes ao tratamento dado à informação ou modelos, para definição

dos níveis de categorização. A síntese é a combinação dos mecanismos anteriores (inspiração, indução e dedução); é a forma mais indicada, pois pode contar para o desenvolvimento da estrutura mais ampla da taxonomia com o auxílio tanto de especialistas quanto da análise dos documentos produzidos e dos modelos e padrões de representação da informação existentes na organização. Isso faz com que se chegue a um resultado mais satisfatório.

Os níveis identificados por qualquer um desses mecanismos são sempre definidos e discutidos antes da disponibilização da taxonomia e, em geral, não necessitam ser modificados.

A taxonomia denominada *beta*, por sua vez, trabalha em um nível mais aprofundado, áreas especializadas do conhecimento. Sendo assim, é necessária a participação de especialistas das áreas de cobertura. Algumas diretrizes que guiaram a equipe da NASA no desenvolvimento da taxonomia *beta* foram:

- Quanto mais especializada for a área de atuação da organização, maior será o nível de especificidade com que as informações serão tratadas.
- Um mesmo conceito pode aparecer em vários lugares da taxonomia, que foi desenvolvida de forma polihierárquica, facilitando assim a recuperação pertinente para o usuário.
- Na estrutura da taxonomia, o conceito e seus relacionamentos devem ser apresentados ao usuário considerando abreviaturas e sinônimos, tornando-a mais acessível.
- Desde o início do desenvolvimento, é importante a adoção de categorias padrões, para que os usuários compartilhem um entendimento comum do que buscam e do tipo de documentos que encontrarão em determinada categoria.

Quanto à construção da taxonomia

Algumas das diretrizes desta prática de elaboração encontram suporte na área da Ciência da Informação, porém não apresentam o devido aprofundamento.

Quanto à determinação das categorias mais amplas (taxonomia *alpha*), são apontados alguns mecanismos que auxiliam esse processo; no entanto, não são discutidos de forma mais detalhada, dificultando a sua execução.

No desenvolvimento da taxonomia *beta*, contar com a participação de especialista é fundamental para a delimitação da especificidade do modelo. Um dos princípios norteadores da proposta da taxonomia *beta* diz que, quanto mais

especializada for a área, mais específicos serão os termos utilizados. Esse processo de especificação faz parte, em metodologias de construção de esquemas de representação do conhecimento, da etapa de análise dos termos, seus relacionamentos semânticos e, juntamente com o controle da diversidade de significação, compõem a categorização.

Na taxonomia *beta*, a especificação está relacionada especialmente ao relacionamento hierárquico, em que um conceito mais específico é aquele que apresenta pelo menos uma característica a mais do conjunto de características do conceito genérico correspondente. Outra peculiaridade desta prática de elaboração refere-se à possibilidade de representação da polihierarquia. Nesse sentido, apesar de não estar descrito de forma bem precisa, pode-se inferir que a prática de elaboração analisada refere-se ao fato de um mesmo conceito poder estar ligado a mais de um termo genérico. Finalmente, a taxonomia *beta* pressupõe o controle de sinônimos e abreviaturas, o que determina o estabelecimento da relação de equivalência como um elemento facilitador do acesso às informações da taxonomia. Apesar de tocar em pontos importantes da construção de taxonomias, como a polihierarquia, os relacionamentos hierárquicos e de equivalência, esta prática de elaboração não indica os procedimentos para realizá-los, de forma que as áreas específicas da organização possam ser representadas eficientemente pela taxonomia.

Quanto à aplicação da taxonomia

Os autores ligam a busca da informação à taxonomia, recomendando que as categorias gerais sejam estáveis, a fim de que o usuário saiba o tipo de documentos que encontrará em determinada categoria quando o resultado da pesquisa lhe é apresentado. É importante que esse aspecto seja construído com os usuários, refletindo a linguagem e fazendo com que a taxonomia atue efetivamente como uma ferramenta de recuperação. A especificação do modelo na construção da taxonomia possibilita a representação específica do assunto do documento, gerando em princípio uma recuperação com alto grau de precisão. No entanto, deve-se prestar atenção quanto à sua utilização sem levar em conta as questões relativas ao critério de revocação, ou seja, aos documentos relevantes não recuperados pelo sistema devido ao alto nível de especificidade adotado na representação do assunto. Segundo Piedade (1977, p.4), “nem sempre a especificidade é uma vantagem, pois

dificulta a recuperação de toda a documentação existente no acervo sobre determinado assunto”. É necessário estabelecer um equilíbrio na busca de um nível de especificidade conveniente.

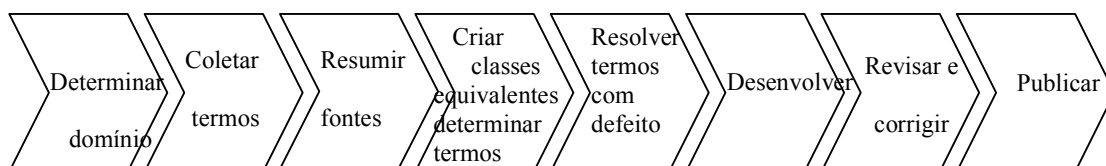
Os autores apontam a forma de construção polihierárquica como uma forma de facilitar a busca, pois um mesmo termo aparecerá em lugares distintos da taxonomia. Essa é também uma preocupação de Woods (2004), que chama de clássica a taxonomia organizada de forma monohierárquica, em que um termo só pode estar ligado a um termo genérico, afirmando que isso não é possível nem desejável no ambiente corporativo.

A recuperação da informação é uma preocupação explícita desta prática de elaboração, especialmente com relação aos resultados da busca. Isso pode ser verificado quando os autores, nas diretrizes, assinalam como necessária a apresentação dos conceitos e seus relacionamentos com os usuários, assim como, a instituição de categorias padrões e a apresentação polihierárquica.

Finalmente, podemos perceber uma interdependência entre o modelo de representação do conhecimento e a representação da informação aplicando a taxonomia, o que contribui para a recuperação eficiente da informação e a organização clara do resultado de uma busca. São processos distintos que se relacionam no momento da categorização.

C) PRÁTICA DE ELABORAÇÃO 3

Kremer; Kolbe e Brenner (2005) estabelecem etapas para o desenvolvimento tanto de uma taxonomia quanto de um glossário, conforme o esquema a seguir:



A primeira etapa consiste em definir qual a abrangência da taxonomia, definindo-a para um setor específico ou para toda a organização. Após essa delimitação, é realizada a coleta de termos com os setores envolvidos, contando

com a participação dos usuários na definição e escolha dos termos. Os autores recomendam, no início do processo, deter-se nos termos mais abrangentes, gerais, contando para isso com o auxílio de fontes, como os sistemas já existentes, dicionários, ou mesmo, analisando os processos organizacionais.

A quarta etapa busca, dentro dos termos mais gerais, agrupar sinônimos e termos relacionados, procurando, já na etapa seguinte, solucionar o que os autores chamam de termos com defeito. Essa definição é usada para termos homônimos, sinônimos ou errôneos; enfim, o objetivo é fazer com que apenas um termo seja usado para determinado conceito e que os conceitos de termos homônimos sejam diferenciados. Na sexta etapa, o objetivo é promover uma classificação lógica, fazendo uso da hierarquia ou do tipo de relação possível entre os termos. Kremer; Kolbe e Brenner (2005) não estabelecem nem discutem o que sejam os outros possíveis tipos de relações entre os termos. Após o estabelecimento e organização dos termos, novamente os usuários auxiliam na avaliação do resultado, testando buscas e analisando a pertinência dos termos alocados. A última etapa é a publicação da taxonomia para uso, procurando também a correção e discussão de possíveis deficiências.

Quanto à construção da taxonomia

A comparação que o autor estabelece entre taxonomia e glossário nos parece um tanto quanto simplista, visto que glossários não apresentam relações semânticas complexas como as taxonomias. No entanto, pode-se considerar que esses conceitos se aproximam na medida em que um glossário contém, em sua estrutura lexicográfica definitória, as informações necessárias para o estabelecimento das relações semânticas de uma taxonomia.

Nessa prática de elaboração, há a indicação da criação de categorias gerais, contando com o auxílio dos usuários e de sistemas ou ferramentas pertinentes. A análise e o controle da diversidade de significações dos termos são desenvolvidos nas etapas quatro e cinco, em que o autor denomina, de forma errônea, termos sinônimos, homônimos como sendo termos com defeito, quando na verdade são as manifestações lingüísticas diferenciadas dos termos. A etapa do desenvolvimento é aquela em que as relações entre termos afins serão estabelecidas. É importante salientar o fato de os autores citarem além das relações hierárquicas, o estabelecimento de outros tipos possíveis de relacionamentos. Cintra et al. (2002)

estabelecem que as relações semânticas são de dois tipos: hierárquicas e não-hierárquicas ou seqüenciais; os relacionamentos seqüenciais são os que apresentam relações funcionais, que denotem atividades ou operações. Caso os autores desta prática de elaboração estejam referindo-se a relações seqüenciais quando se referem a 'outros tipos de possíveis relacionamentos', pode-se deduzir que estes são relacionamentos semânticos expressivos no contexto das organizações. No entanto, apesar de a prática de elaboração deixar explícita a possibilidade de as taxonomias irem além das relações hierárquicas, não são indicados claramente quais seriam esses tipos possíveis de relacionamentos. Concluindo, a prática de elaboração apresenta-se de forma superficial, dificultando assim a construção de uma taxonomia e suas relações semânticas.

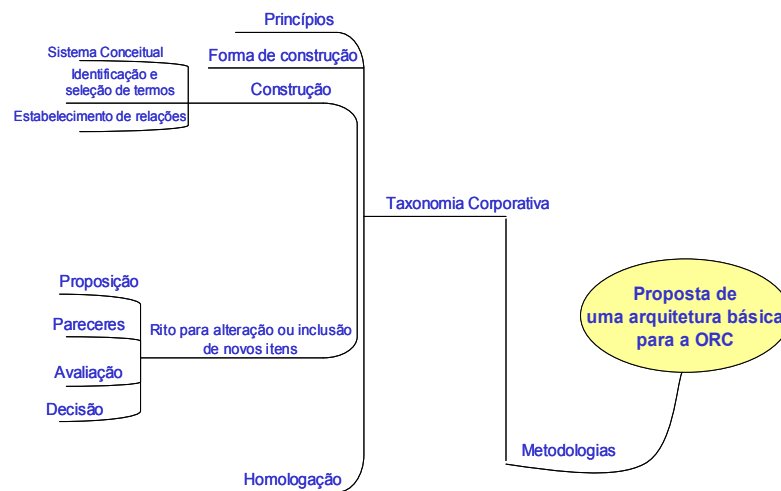
Quanto à aplicação da taxonomia

As últimas etapas mostram a preocupação com a avaliação dos resultados da taxonomia, buscando a revisão junto aos usuários. A indicação para que haja um teste de busca juntamente com a avaliação dos termos demonstra que, após o processo de representação, a taxonomia assume o papel de ferramenta de recuperação, auxiliando na avaliação dos termos pelas buscas e resultados obtidos. A maneira como se procede à avaliação não é explicitada na prática de elaboração, assim como o fato de a taxonomia organizar as informações com objetivo de recuperação; todavia, está implícito nos processos mencionados o fato de ela servir de instrumento para esse fim.

D) PRÁTICA DE ELABORAÇÃO 4

Yamaoka (2005) desenvolveu uma proposta de arquitetura para a organização e representação do conhecimento (ORC). Iremos aqui nos deter nas etapas do desenvolvimento da taxonomia corporativa.

O desenvolvimento da taxonomia corporativa está dividido em cinco etapas gerais: princípios, forma de construção, construção, rito para alteração ou inclusão de novos itens e a homologação, conforme figura abaixo.



Os princípios se referem aos objetivos da criação da taxonomia na organização, que pode ser tanto para auxiliar os usuários na busca da informação, quanto para diagnosticar o comportamento de busca e/ou necessidades informacionais. O autor apresenta alguns princípios que podem variar, dependendo do ambiente onde a taxonomia for desenvolvida. Na etapa 'forma de construção', é apontada a necessidade de participação de especialistas e usuários de diferentes áreas envolvidas no processo.

A parte central da metodologia para o desenvolvimento da taxonomia corporativa é a etapa da construção. Nela, é necessário, segundo o autor, delimitar um sistema conceitual condizente com os princípios propostos, isto é, que atendam às necessidades informacionais de seus usuários. Para apoiar esta etapa, Yamaoka (2005) propõe uma base de dados terminológica orientada para as etapas de identificação e seleção dos termos e o estabelecimento das relações conceituais entre eles. O conteúdo da organização - processos informacionais já existentes e seus usuários - também auxilia no processo de identificação e seleção dos termos.

Os elementos da taxonomia corporativa relacionados à base de dados terminológica são:

- O controle terminológico das designações dos termos da taxonomia corporativa.
- O estabelecimento das relações conceituais nas taxonomias corporativas.
- O desenvolvimento de outras taxonomias que forem necessárias para áreas ou objetivos específicos da empresa.
- A escolha correta das palavras-chave para serem utilizadas para descrever os recursos de informação, no elemento de metadados "assuntos". (YAMAOKA, 2005)

O desenvolvimento de uma taxonomia, por ser um processo dinâmico, necessita estar em constante revisão. Nesse sentido, Yamaoka (2005) estabelece algumas etapas para a manutenção contínua de uma taxonomia. São elas:

- Proposição: a solicitação da alteração é encaminhada para uma equipe, chamada escritório de ORC.
- Pareceres: Essa equipe encaminha para os setores envolvidos a solicitação, afim de que dêem um parecer a respeito da coerência e impacto da mudança.
- Avaliação: Com base nesses pareceres, é realizada uma avaliação.
- Decisão: Após essa avaliação, a solicitação é ou não aceita.

A última etapa da taxonomia é a homologação, que, segundo o autor, “estabelece um processo de verificação da adequação da taxonomia em relação aos usuários e ao conteúdo que representa”.

Quanto à construção da taxonomia

A prática de elaboração não apresenta a forma pela qual os termos serão selecionados e quais categorias serão construídas. Para a etapa da construção, o autor remete a um outro elemento da proposta, a base de dados terminológica. Entretanto, nesta etapa são apresentadas as diretrizes, mas não a forma como a construção efetivamente acontecerá. Na proposta da base de dados terminológica, é indicado o controle terminológico dos termos, o estabelecimento das relações, o desenvolvimento de outras taxonomias para objetivos específicos e a escolha das palavras-chave que irão descrever os recursos informacionais. No entanto, não há um aprofundamento teórico e metodológico sobre a maneira como esses objetivos serão atingidos, principalmente no que se refere ao estabelecimento das relações conceituais.

O autor menciona a dinâmica do processo de desenvolvimento de taxonomias como justificativa para uma constante revisão da terminologia e para isso apresenta algumas etapas, bem estabelecidas, para a inclusão de termos. Essas etapas são apresentadas no ‘Rito para alteração ou inclusão de novos itens’, que estabelece que, para a inclusão de novos termos, são necessários uma proposição e o parecer da equipe responsável a respeito da coerência e impacto da mudança. Com base nesse parecer, é feita uma avaliação para só então ser tomada a decisão de aceitar ou não a inclusão do novo item.

A prática de elaboração não apresenta as etapas de desenvolvimento das construções semânticas, como citado, remetendo para um segundo elemento da proposta que trata desta etapa - a base de dados terminológica.

Quanto à aplicação da taxonomia

A primeira etapa desta prática de elaboração, a definição dos princípios, indica a taxonomia como uma ferramenta auxiliar na busca e no diagnóstico do comportamento de busca. Como ferramenta de apoio à busca, esta etapa deixa claro o papel fundamental da taxonomia no momento da pesquisa e organização da informação recuperada. Quanto a ser um instrumento de suporte ao diagnóstico do comportamento de busca, uma análise desse comportamento poderá evidenciar a eficiência da taxonomia no processo de recuperação da informação, oferecendo informações importantes para o aperfeiçoamento da prática de elaboração. Na realização do diagnóstico do comportamento de busca, é possível avaliar de que forma a taxonomia está auxiliando o usuário na construção de estratégias ou mesmo se as relações construídas fazem sentido para aquele ambiente.

Esses princípios mostram igualmente, como apontam Terra e Gordon (2002), que as relações entre os termos e a própria organização da taxonomia têm a função de oferecer ao usuário novas possibilidades no que tange à recuperação, ampliando seu campo de percepção. Entretanto, a prática de elaboração não apresenta a forma pela qual o diagnóstico de busca deve ser realizado.

Na prática de elaboração, é indicada ainda a necessidade da homologação, que se refere à adequação da representação da informação às necessidades informacionais dos usuários. Esse processo pode ser balizado pelo diagnóstico do comportamento de busca.

Enfim, fica evidente nesta prática de elaboração de taxonomia a ligação entre a representação e a organização da informação com apoio da taxonomia no processo de recuperação.

E) PRÁTICA DE ELABORAÇÃO 5

O Montague Institute (2001) divide a taxonomia em duas partes: a estrutura e a aplicação. A estrutura consiste nas categorias ou termos e as relações entre eles;

a aplicação se refere às ferramentas de navegação que auxiliam o usuário a localizar a informação. Uma parte é dependente da outra.

A taxonomia tem início com a definição dos termos padrões que descrevem, de forma sumarizada, as informações constituintes da estrutura. A segunda e terceira etapas consistem em construir as relações entre os termos. A estrutura da taxonomia é composta, portanto, pelos seguintes elementos:

- Lista de termos padrões.
- Relações hierárquicas.
- Referências cruzadas.

Na prática de elaboração, relata-se que, após a construção dos relacionamentos entre termos, é necessário conectá-los com os recursos disponíveis (documentos, pessoas, e-mails, *Web sites*, etc.), momento da aplicação da taxonomia.

Quanto à construção da taxonomia

Nesta prática de elaboração há uma preocupação do autor com o processo de construção conceitual da taxonomia. Em virtude disso, ele a divide em duas partes, estrutura e aplicação, sendo a primeira referente a esse processo. A descrição sumarizada da estrutura da taxonomia pode ser comparada ao estabelecimento das categorias gerais. Na etapa da construção das relações semânticas, há a indicação do estabelecimento das relações cruzadas 'Ver' e 'Ver também', além das relações hierárquicas. A referência 'Ver' faz parte da relação de equivalência e é definida no momento do controle da diversidade terminológica existente para um mesmo conceito, utilizado em casos de sinônimas. Já a referência 'Ver também' é considerada uma relação associativa, que pode ser de vários tipos, conforme Cintra et al. (2002). A prática de elaboração já sinaliza com a possibilidade da existência de relações além das hierárquicas no processo de construção. Nesse trabalho, percebemos uma preocupação com o processo de construção da categorização; notamos, entretanto, a falha na falta de um detalhamento dos procedimentos para o desenvolvimento da taxonomia. O que se tem é somente uma lista das etapas desse processo. Outro ponto observado se refere à omissão da maneira como se procederá a coleta e análise dos termos que constituirão a taxonomia.

Quanto à aplicação da taxonomia

Há uma indicação de que, após a construção dos relacionamentos semânticos, ocorra a conexão dos termos com os recursos disponíveis. O autor cita como exemplo desses recursos os documentos, pessoas, e-mails e outros considerados importantes para o grupo. Podemos perceber assim a ligação entre a construção da taxonomia e a recuperação da informação. Terra et al. (2005); Dutra e Bush (2003) e o Delphi Group (2002) reconhecem a taxonomia como um instrumento de recuperação da informação pelos usuários de uma organização, ressaltando a importância da integração eficiente entre os dois processos: a construção e a aplicação da taxonomia na recuperação. A divisão da taxonomia em estrutura e aplicação deixa clara a diferença entre os aspectos de representação do conhecimento, representação da informação e organização da informação recuperada. Conforme esta prática de elaboração, a aplicação objetiva localizar a informação no ambiente do portal. Sendo assim, o processo de estruturação serve de base para uma posterior aplicação, facilitando e viabilizando as buscas de informações.

F) PRÁTICA DE ELABORAÇÃO 6

A sexta prática de elaboração de taxonomia é apresentada pela empresa de consultoria Kcurve. A autora, Marcia Morante, trabalha com desenvolvimento de taxonomias para organizações. Morante (2003) estabelece cinco etapas para o desenvolvimento de uma taxonomia:

1) Determinar o domínio e o alcance

- Conhecer o ambiente a fim de avaliar:
- Conteúdo: alcance, nível de especificidade, formatos e tipos de documentos, integração dos dados, etc.
- Tecnologias e integração de funções: mecanismos de busca, sistema para o gerenciamento de conteúdo, existência de portal, etc.
- Recursos de informação: existência de taxonomias e/ou esquemas de navegação.

- Necessidades dos usuários: quem são os usuários, quais os conteúdos necessários, que tarefas executam e de que informações precisam, se existem variações de vocabulário dentro da mesma organização.

2) Revisão das autoridades de assunto

- Verificar a existência de vocabulários de associações de indústrias, instituições acadêmicas, que possam ser adaptados às necessidades da organização.
- Consultar especialistas da área.
- Incorporar os termos padrões.

3) Reunir os termos

- Realizar auditoria de conteúdo e registros: tipos e formatos, palavras-chave, metadados.
- Explorar outros recursos: tesouros e esquemas de navegação.
- Modificar um vocabulário existente.

4) Organizar os termos

- É importante a participação do usuário nesta etapa.
- Identificar e nomear altos níveis de conceitos, que seriam as facetas.
- Procurar similaridade entre os termos previamente localizados.
- Determinar listas e/ou hierarquias.
- Definir termos preferenciais e termos relacionados.
- Ligar sinônimos e variações.

5) Validação e Teste

- Realizar comparações entre as necessidades dos usuários e os níveis gerais de assuntos estabelecidos na taxonomia, assim como o nível de especificação destes.
- Testar usabilidade.
- Rever todas as etapas anteriores.

Quanto à construção da taxonomia

A etapa 1 refere-se ao conhecimento do ambiente, envolvendo verificações relacionadas à construção e aplicação da taxonomia. Constitui-se em uma etapa de reconhecimento do ambiente. Ressalta-se a preocupação nesta prática de elaboração, conforme referenciado nas práticas de elaboração 1 e 2, em verificar a existência de outros vocabulários controlados ou esquemas de classificação. É

importante conhecer o que já foi construído anteriormente e as causas pelas quais não podem ser usados. As etapas 2 e 3 estão direcionadas à análise dos termos selecionados, tanto nos sistemas existentes como junto aos usuários.

Com base nessas etapas, realiza-se a organização dos termos na etapa 4. O primeiro procedimento da organização terminológica é a identificação e nomeação dos altos níveis de conceitos, o que a autora da prática de elaboração chama de facetas. De acordo com Langridge (1973 *apud* PIEDADE, 1977), “usa-se Categoria quando se refere à estrutura geral de um sistema de classificação; faceta quando se refere às manifestações das categorias nas diferentes classes; e o princípio usado para definir uma faceta é conhecido como característica da divisão”. Sendo assim, uma faceta somente pode existir se for aplicada uma característica como princípio de divisão da categoria. Como a prática de elaboração não esclarece se utiliza esse procedimento para definir as facetas, a denominação categoria geral parece-nos mais apropriada.

Logo após o estabelecimento das categorias, é realizada a análise dos termos a fim de organizá-los no que a autora chama de listas e/ou hierarquias. Não fica claro na prática de elaboração se existe diferença entre os conceitos de lista e hierarquia; existindo, também não esclarece qual seria essa diferença. Da forma como está descrito este item na prática de elaboração analisada, parece que hierarquias também podem ser consideradas listas, o que pode causar uma ambigüidade no momento da construção da taxonomia.

No momento da organização, a prática de elaboração instrui o elaborador a definir termos relacionados e preferenciais. Isso poderia ser interpretado como a definição de relações seqüenciais e de equivalência, apesar de não haver essa especificação na prática de elaboração. Em seguida, a prática de elaboração determina a ligação de sinônimos e variações, contemplando somente o grupo de termos preferenciais definido anteriormente. Fica a dúvida quanto ao que fazer com os termos relacionados definidos. Essas inconsistências, no entendimento da prática de elaboração, podem resultar em uma taxonomia pouco eficiente.

Quanto à aplicação da taxonomia

A etapa de validação e teste da taxonomia nesta prática de elaboração apresenta uma preocupação com a recuperação da informação. A autora aponta o fato de que as categorias gerais e suas especificações sejam condizentes às

necessidades dos usuários. O atendimento dessas necessidades pode ser avaliado no momento do uso da taxonomia, na recuperação, o que é denominado teste de usabilidade.

O quadro 7, a seguir, apresenta a síntese da análise das práticas de elaboração de taxonomia. O objetivo é sistematizar os dados de maneira a facilitar a visualização de que as práticas de elaboração analisadas não apresentam todas as etapas para a construção de uma taxonomia corporativa com o devido embasamento teórico e metodológico.

Etapas de Construção e Aplicação da Taxonomia	PRÁTICAS DE ELABORAÇÃO (PE)					
	PE 1	PE 2	PE 3	PE 4	PE 5	PE 6
Estabelecimento das categorias gerais	Estabelece as categorias por meio do que chama 'auditoria do conhecimento', com a participação dos usuários.	Apresenta mecanismos para estabelecimento das categorias gerais.	Indica o auxílio do usuário e de sistemas ou ferramentas pertinentes no processo.	Não explicita a etapa.	Indica a descrição sumarizada da estrutura.	Identifica e nomeia os altos níveis de conceitos, sem a explicitação da metodologia.
Coleta dos termos	Indica o uso de informações consideradas indispensáveis para o grupo e termos que as definam.	Estabelece o uso de exemplares da documentação produzida e de vocabulários controlados.	Realiza junto aos setores e processos organizacionais, além de fontes disponíveis.	Não explicita a etapa.	Não explicita a etapa.	Indica o uso de tesouros, vocabulários especializados e especialistas.
Análise dos termos selecionados	Não apresenta critérios de análise.	Especifica os termos com auxílio de especialistas.	Junto aos setores/processos organizacionais e fontes disponíveis.	Não apresenta critérios de análise.	Não apresenta critérios de análise.	Realiza por meio da busca da similaridade entre os termos.
Controle da diversidade de significação	Busca definir os diferentes significados que um mesmo termo possa ter no grupo (polissemia).	Não estabelece o desenvolvimento da etapa de forma clara.	Agrupar sinônimos e termos relacionados.	Indica o controle terminológico, sem explicitação.	Denota a identificação dos termos sinônimos, pela relação 'Ver'.	Estabelece os termos preferenciais, ligar sinônimos e variações.
Construção dos relacionamentos semânticos	Não explicita a forma na qual os termos serão organizados no modelo.	Denota o relacionamento hierárquico e polihierárquico.	Denota o relacionamento hierárquico e 'outros tipos de relações possíveis'.	Indica a construção das relações semânticas, sem explicitação.	Estabelece o uso das relações cruzadas 'Ver' e 'Ver também', além das hierárquicas.	Indica a definição de termos relacionados e preferenciais.
Representação da informação (classificação ou indexação)	Cita a realização de um teste de busca.	Busca a representação específica do assunto do documento.	Indica a avaliação dos resultados, revisando junto aos usuários.	Salienta a necessidade de homologação dos termos.	Não há indicações explícitas.	Não há indicações explícitas.
Organização da informação recuperada	Não explicita a etapa.	Salienta o uso de categorias estáveis no momento da recuperação.	Não explicita a etapa.	Indica a taxonomia como ferramenta de apoio às buscas, sem explicitação.	Aponta a conexão dos termos com recursos disponíveis, para uma possível recuperação.	Apresenta a informação organizada em categorias relevantes aos usuários.

Quadro 7: Síntese da análise das práticas de elaboração de taxonomia

Fonte: Do autor

5 RECOMENDAÇÕES PARA O DESENVOLVIMENTO DE UMA TAXONOMIA EM PORTAIS CORPORATIVOS

A taxonomia vista como um sistema de classificação, definido por Langridge (1987) como “um mapa completo de qualquer área do conhecimento, mostrando todos os seus conceitos e suas relações [...]”, é um objeto de estudo antigo da Documentação. Porém, no contexto dos ambientes organizacionais, sua aplicação é relativamente recente. Nesse meio, ela surge como uma forma de melhoria da organização da informação em portais corporativos, possibilitando o tratamento adequado dos documentos com objetivo de recuperá-los.

Apesar de se respaldar em algumas etapas de desenvolvimento de tesouros, entendemos a taxonomia como um modelo diferenciado, que tem como proposição atuar em ambientes organizacionais, tratando de áreas específicas do conhecimento como apoio à gestão, tanto da informação quanto do conhecimento. Além disso, a sua aplicação é sempre direcionada a portais, tendo como característica o uso dos recursos inerentes a esse ambiente. A taxonomia corporativa não requer a aplicação de critérios rígidos no desenvolvimento dos relacionamentos semânticos, o foco central é ter significação aos grupos de usuários para o qual foi desenvolvido.

Dessa maneira, enquanto os portais corporativos exercem a função de armazenamento e difusão da informação produzida pelas empresas, a taxonomia atua como instrumento mediador entre aquele que registra o conhecimento em forma de informação e o que busca e recebe a informação procurada. Nesse sentido, pode-se dizer que a taxonomia corporativa é um índice dos vários documentos dispersos na empresa e contribui na obtenção de altos níveis de relevância no processo de recuperação. Assim, ela está estreitamente ligada às teorias e métodos elaborados pelas áreas de Classificação, Indexação, Lingüística e Documentação, nos quais deve apoiar-se. Respaldado teórica e metodologicamente por essas áreas, expõem-se neste capítulo algumas recomendações, tomando por base a análise das práticas de elaboração identificadas na literatura sob a ótica da funcionalidade categorização. A organização do capítulo segue a sistemática do capítulo relativo à análise dos dados. Assim, a primeira parte trata da construção da taxonomia, enquanto a segunda é dedicada à aplicação da taxonomia na representação da informação e

organização dos documentos recuperados. Apesar das recomendações objetivarem o desenvolvimento de taxonomias no ambiente corporativo, suas indicações podem ser aplicadas a qualquer ambiente, realizando as devidas adaptações.

5.1 CONSTRUÇÃO DA TAXONOMIA

A construção da taxonomia corporativa refere-se ao processo de representação do conhecimento existente no ambiente corporativo. Seu desenvolvimento deve estar orientado pelas cinco etapas descritas a seguir.

1) Estabelecimento das categorias gerais

O estabelecimento das categorias como primeira etapa de construção da taxonomia deve seguir tanto uma abordagem dedutiva como indutiva.

Segundo Vickery (1960), no método dedutivo estabelecem-se as categorias gerais em um trabalho de equipe, entre especialistas, profissionais da informação e usuários, de acordo com as necessidades percebidas por esse grupo. A esse grupo cabe prever todas as possíveis categorias que representem o universo informacional da organização, constituindo assim a estrutura geral da taxonomia. Essa orientação ao trabalho cooperativo foi observada nas práticas de elaboração 1, 2 e 3.

A literatura considera a Teoria da Classificação Facetada, definida por Ranganathan, como a principal representante deste método. Nela, são definidas as categorias fundamentais as quais demarcam a primeira classificação de assuntos dentro de um grande universo. Campos (2001) reconhece que “são elas que fornecem a visão de conjunto dos agrupamentos que ocorrem na estrutura, possibilitando assim o entendimento global da área”.

Conforme visto no item 2.3.1, Ranganathan (1962) instituiu cinco categorias fundamentais:

Personalidade – É considerada indefinível, quando um termo não se adapta a nenhuma outra categoria, é entendida como uma manifestação desta.

Matéria – Materiais constitutivos de um objeto. Ex: mesa/Matéria: madeira.

Energia – Processos, operações, ações, técnicas, métodos, fenômenos.

Espaço – Lugares.

Tempo – Períodos.

Alguns autores apontam a insuficiência dessa lista e sugerem sua ampliação. É o caso, por exemplo, do elenco de categorias proposto pelo *Classification Research Group* – CRG (1955) e exposto abaixo.

Tipos de produto final

- Partes – divisões do produto final
- Materiais – constituintes do produto final
- Propriedades – qualidades do produto final
- Processos – Ações ou reações das próprias coisas
- Operações – ações sofridas pelas coisas
- Agentes – responsáveis pelas ações
- Espaço – lugar
- Tempo - época
- Forma de apresentação – tipo de documento.

Segundo Piedade (1977, p.13), nas categorias propostas pelo CRG, “a razão de ser do estudo do assunto a classificar é o seu produto final”. No caso de um setor de desenvolvimento de sistemas, por exemplo, o produto final é o software; suas partes são os diferentes aplicativos; seu material é o disco óptico; a propriedade, como qualidades do produto final, é, no caso do software, a portabilidade; os processos são a instalação dos programas, o gerenciamento de dados; as operações são planejar, projetar e desenhar; o agente é o programador e a forma de apresentação é o *cd-rom*, por exemplo.

Esses dois exemplos trazidos da literatura da área de Biblioteconomia podem servir como guia para o mapeamento das categorias existentes no ambiente corporativo. Observamos, no entanto, que, devido à natureza do objeto desse tipo de taxonomia, categorias que representam ‘o que é produzido’ (objeto), ‘quem produz’ (agente) e ‘como é produzido’ (processo) são imprescindíveis. Nesse sentido, o esquema proposto pelo CRG parece ser mais apropriado.

Juntamente com o método dedutivo, as categorias podem ser estabelecidas pela abordagem indutiva que busca defini-las de acordo com as diferentes fontes da área. Vickery (1960) diz que essas categorias podem ser determinadas pela consulta à literatura especializada da área, como dicionários, tesouros, artigos, relatórios técnicos, etc. É importante que se tome como base classificações ou estruturas

informativas existentes e se possa validá-las por um grupo de especialistas ou representantes das diferentes áreas da organização. Como orientação ao método indutivo, citamos as orientações apresentadas pelas práticas de elaboração 1, 2 e 3 que observam a importância de:

- a) Questionar-se sobre que categorias são importantes para o grupo (trabalho cooperativo denominado 'autoria do conhecimento' conforme prática de elaboração 1).
- b) Pesquisar sobre as formas preferenciais de navegação e classificação em sites utilizadas pelo grupo.
- c) Conhecer as maneiras pessoais de organização de cada integrante do grupo.

Observamos, finalmente, a necessidade de denominar corretamente as categorias, notando aspectos de variação conceitual e terminológica.

2) Coleta dos termos

Na compilação dos termos, a consulta à literatura especializada ou classificações existentes na organização deve sempre ser utilizada como recurso auxiliar. De acordo com Ramsden (1974 *apud* PIEDADE, 1977), como os termos precisam fazer sentido ao grupo de usuários que irá utilizá-los, a literatura especializada serve como validação dos termos selecionados. Sendo assim, dois princípios devem ser levados em consideração no momento da coleta de termos: a garantia literária e a garantia de uso.

Lancaster (1987, p.21), ao tratar sobre o tema garantia literária, registra que “[...] um termo se justifica apenas se ocorre dentro da literatura de um determinado assunto com algum grau de frequência”. Esse princípio pode ser aplicado quando se coletam termos de fontes de referência significativas da área, que pode ser um tesouro ou publicações de resumos, por exemplo. Com os sistemas automatizados e as bases de dados, esse trabalho torna-se mais rápido e eficiente.

Quanto ao conceito de garantia de uso, os termos coletados precisam representar as necessidades dos usuários. Lancaster (1987) afirma que “alguns termos podem ocorrer na literatura, mas serem de pouco interesse para um determinado grupo de usuários”. Daí a importância da participação do especialista

na coleta dos termos, seja apresentando seus assuntos de interesse ou esquemas de organização da informação, seja por meio da análise de registros de buscas já realizadas.

É preciso haver um equilíbrio entre esses dois princípios, pois é fundamental o atendimento às necessidades dos usuários.

A prática de elaboração 2 aponta a necessidade de recolher exemplares da documentação produzida e de vocabulários controlados existentes na organização a fim de direcionar o desenvolvimento dessa etapa. Como procedimento de garantia de uso, ressaltamos o proposto pela prática de elaboração 4 denominado “Rito para alteração ou inclusão de novos termos”. Por meio dele, os termos são sugeridos com base em um parecer que é avaliado pela equipe de construção da taxonomia.

3) Análise dos termos selecionados

Após a coleta dos termos, é necessário padronizá-los, para que todos os usuários ‘falem a mesma língua’. Desse modo, são estabelecidas algumas diretrizes, baseadas em Cavalcanti (1978).

- **Singular e plural:** Primeiramente, verifica-se como o termo é utilizado no ambiente e se o significado muda no caso de uso do plural ou singular, como, por exemplo, Memória = faculdade de reter as idéias e Memórias = registros de alguém que conta sua vida. Em geral, há a indicação do uso do singular nos seguintes casos: materiais (areia, cimento); propriedades específicas (condutividade); processo (seleção); nomes próprios (Unesco); disciplinas (matemática, gestão de pessoas). Nos casos de termos genéricos que designam classes de objetos, relações humanas, móveis, contratos de trabalho, orienta-se o uso do plural.

- **Abreviaturas e siglas:** Usar somente quando facilmente reconhecidas ou quando o conceito é mais conhecido pela sigla, como por exemplo, PIB, ICMS, CPMF. Nesses casos, sempre construir remissivas para o nome em extenso, mantendo a consistência.

- **Empréstimos lingüísticos:** A tradução para o português de termos em outras línguas dependerá da incorporação que esse termo sofreu na língua de especialidade, como, por exemplo, *site* ou sítio, *e-mail* ou correio eletrônico, *call center* ou central de atendimento. Alguns termos, em outras línguas, não

apresentam correspondentes na língua portuguesa e são utilizados na forma original, como, *royalty*, *habeas corpus*, *software*.

- **Nomes populares e nomes científicos:** A escolha de um ou de outro dependerá da área de atuação da organização e da preferência dos usuários.

- **Gíria ou jargão:** Dependerá das preferências lingüísticas dos usuários, refletindo a manifestação lingüística mais freqüente no ambiente.

Excluindo as questões relativas a singular ou plural, deve-se utilizar a relação de equivalência (ver etapa 5) para o controle das denominações de natureza sinonímica.

4) Controle da diversidade de significação

Em ambientes em que o objetivo seja o compartilhamento de uma mesma linguagem, é necessário neutralizar a ambigüidade, entendida por Cintra et al.(2002) “[...] como a possibilidade de uma comunicação lingüística prestar-se a mais de uma interpretação”. A ambigüidade se impõe por meio da polissemia e da homonímia (CINTRA et al., 2002). A polissemia diz respeito à diversidade de significação de um mesmo termo; a homonímia refere-se a palavras com identidade de sons ou de forma, mas com significados distintos. Essas ambigüidades podem ser controladas por meio da contextualização do significado do termo. Essa contextualização é registrada entre parênteses como ilustrado nos exemplos a seguir.

- companhia (empresa); companhia (pessoa)

- capital (dinheiro); capital (sede do país)

Para neutralizá-las, pode-se também, se necessário, elaborar “[...] redes relacionais e [estabelecer] definições e notas de escopo, sempre que as redes se mostrarem insuficientes para a interpretação unívoca da significação” (CINTRA et al., 2002). Cavalcanti (1978) cita como exemplo o termo composto ‘Organização da fábrica’ que, para o sentido ser entendido, é necessário instituir uma nota explicativa (NE), como exemplificado abaixo:

NE: esquema de trabalho de um fábrica.

A elaboração das redes relacionais, auxiliares no controle da polissemia e da homonímia, será detalhada na próxima etapa.

Apesar de algumas das práticas de elaboração analisadas nesta pesquisa apontarem a problemática da ambigüidade, eles não aprofundam esta questão; portanto, não apresentam contribuição que possa ser utilizada como recomendação.

5) Construção dos relacionamentos semânticos

Nesta etapa, levam-se em consideração as relações semânticas de equivalência, hierárquicas e não-hierárquicas assim como a polihierarquia.

As relações de equivalência se estabelecem entre termos sinônimos ou quase-sinônimos, além de alguns itens apresentados na etapa 3. A análise dos termos sinônimos precisa corresponder às necessidades e preferências do grupo de usuário. Seu controle deve ser realizado, permitindo a padronização e conseqüente aumento da qualidade da taxonomia. É importante que essa 'negociação' ocorra em parceria entre aquele que desenvolve a taxonomia e os usuários do sistema. Recomenda-se a utilização dos símbolos USE e UP (usado para). O USE tem a função de indicar o termo preferencial, enquanto que o UP tem o papel de indicar os termos proibidos, ou seja, aqueles que não podem ser utilizados para representar um assunto. A seguir, ilustra-se a utilização desses símbolos para o controle dos sinônimos EMPRESA e ORGANIZAÇÃO.

EMPRESA

USE ORGANIZAÇÃO

ORGANIZAÇÃO

UP EMPRESA

As relações hierárquicas podem ser do tipo gênero/espécie ou todo/parte.

Relação gênero/espécie: exprime os graus de superordenação e subordinação entre os conceitos. O termo superordenado representa o termo genérico (TG), enquanto o termo subordinado representa o termo específico (TE).

MOVIMENTAÇÃO FINANCEIRA

TE APLICAÇÃO

APLICAÇÃO

TG MOVIMENTAÇÃO FINANCEIRA

O termo específico representa um conceito subordinado que possui as mesmas características do conceito geral e mais alguma que o diferencia de outro conceito

específico subordinado ao mesmo conceito geral. Ele é considerado um tipo do termo genérico. No exemplo citado, APLICAÇÃO é um tipo de MOVIMENTAÇÃO FINANCEIRA.

Relação todo/parte: ocorre entre um objeto considerado como um todo e suas partes. Utiliza os mesmos símbolos da relação gênero/espécie. Exemplo:

COMPUTADOR

TE HARDWARE

HARDWARE

TG COMPUTADOR

Lancaster (1987, p.26) acrescenta que

[...] certas relações partitivas são freqüentemente tratadas como se fossem genéricas. Partes do corpo e nomes geográficos são os melhores exemplos. Faz sentido fazer de 'aorta' um TE subordinado a 'coração', embora ela seja uma parte e não um tipo de coração, o mesmo valendo para localização geográfica.

O mesmo acontece em relação a campos de estudo e estruturas sociais que são relações partitivas tratadas como genéricas.

A **polihierarquia** também deve ser considerada no estabelecimento das relações semânticas. Nela, um mesmo termo específico pode estar subordinado a mais de um termo genérico. Apenas a prática de elaboração 2 demonstra a preocupação com a representação dessa situação de significação. As relações de polihierarquia são freqüentes no ambiente organizacional. O termo 'matéria-prima' é um exemplo. Na área financeira, este conceito é considerado como um item de custo; na área de produção é um tipo de insumo. O termo 'matéria-prima' aparecerá ligado a dois termos genéricos. O que define a sua relação são as ligações semânticas estabelecidas na estrutura conceitual na qual está inserido.

As **relações não-hierárquicas ou seqüenciais** são muito freqüentes em ambientes organizacionais por representarem uma diversidade muito maior de relacionamentos semânticos. De acordo com Cintra et al. (2002), relações não-hierárquicas ou seqüenciais são as que apresentam entre si contigüidade espacial ou temporal. Algumas tipologias de relações seqüenciais são apresentadas no capítulo 2 (item 2.4). No ambiente organizacional as que aparecem com maior freqüência são:

os conceitos no nível de especificidade²⁶ necessário, viabilizará a representação específica dos assuntos dos documentos. Esse fato provavelmente aumentará o nível de precisão²⁷ na recuperação da informação. Por outro lado, a adoção do relacionamento não-hierárquico auxiliará a adoção do princípio de exaustividade, ou seja, a representação de todos os assuntos presentes em um documento, podendo colaborar com os níveis de revocação²⁸ na recuperação da informação. Além disso, quando é utilizada uma taxonomia, os documentos têm maior probabilidade de serem indexados da mesma forma, atingindo assim a uniformidade no tratamento da informação e de sua recuperação.

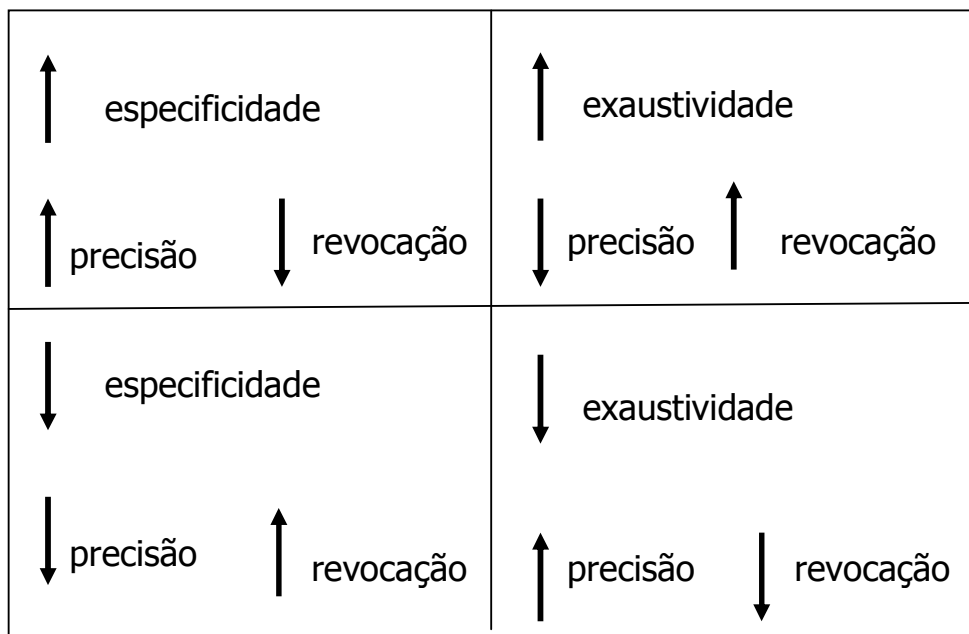
No Quadro 7, são apresentados os princípios de especificidade e exaustividade em relação aos critérios de revocação e precisão de um sistema de recuperação da informação. Percebe-se que, quanto maior o grau de especificidade, maior será a precisão e menor a revocação na recuperação. Na medida em que se diminui a especificidade, menor será a precisão, havendo uma probabilidade de aumento da revocação. A adoção do princípio de exaustividade atua de forma oposta. Quanto maior a exaustividade, maior a revocação e menor a precisão; quanto menor a exaustividade, menor a precisão e maior a revocação.

Sendo a opção pela especificidade e pela exaustividade diretamente dependente da correta construção de uma taxonomia, podemos deduzir que a qualidade da taxonomia influencia a qualidade da recuperação da informação.

²⁶ “Exatidão com que os descritores utilizados representam o conteúdo temático dos documentos”. (PIEADADE, 1977, p.4)

²⁷ “Relação entre os documentos relevantes recuperados e o número total de documentos recuperados”. Ibid, p.3

²⁸ “Relação entre os documentos relevantes recuperados e o número total de documentos relevantes sabidamente existentes na coleção”. Ibid, p.3



Quadro 8: Revocação e precisão em relação a especificidade e exaustividade
Fonte: Café (2006)

Em relação à **organização da informação** recuperada do acervo, a taxonomia proporciona a categorização dos resultados de um busca, facilitando a identificação das relações entre os assuntos. O mecanismo de busca do portal deve apresentar a saída dos resultados organizados nas categorias previamente definidas. Dessa forma, o usuário perceberá outros assuntos relacionados ao tema procurado que, muitas vezes, não lhe haviam ocorrido. Essa preocupação está presente nas práticas de elaboração 1 e 2 analisados no corpus da pesquisa. As categorias gerais estabelecidas na etapa 1 para a organização dos termos são as mesmas utilizadas no momento da recuperação. A prática de elaboração 5 apresenta a necessidade de conexão entre a documentação do portal e as categorias da taxonomia no momento da aplicação, imprescindível para que a taxonomia auxilie no processo de recuperação, com link direto aos recursos informacionais.

Conforme salienta Dutra e Busch (2003), a taxonomia precisa servir de referência para o entendimento da área do conhecimento que se dispõe a representar.

6 CONCLUSÃO

A seguir são apresentadas as conclusões advindas da pesquisa e propostas para futuros trabalhos.

No contexto organizacional, a informação, propulsora da criação do conhecimento, constitui-se em diferencial competitivo pela capacidade de gerar inovações, aperfeiçoando processos e produtos. Devido a esse papel estratégico, as organizações desenvolvem ações no sentido de qualificar o processo de armazenamento, busca e recuperação de informações.

Nesse sentido, são desenvolvidos portais corporativos que buscam a criação de um ponto de acesso comum ao universo informacional da organização, objetivando sua integração e o compartilhamento das informações. Para atingir esse objetivo, é preciso haver um modelo de representação do conhecimento. A taxonomia é um tipo desse modelo utilizado em portais corporativos, garantindo a categorização e organização dos recursos informacionais nesse ambiente.

Esta pesquisa buscou propor princípios básicos para o processo de construção de taxonomias em portais corporativos, no sentido de apresentar uma discussão teórica mais aprofundada acerca de suas etapas. A proposição das etapas de construção da taxonomia partiu da análise de seis práticas de elaboração localizadas na literatura da área de Ciência da Informação, juntamente com os objetivos da funcionalidade 'categorização', apresentada na literatura da área de portais corporativos e relacionada ao processo de representação do conhecimento.

Para o alcance do objetivo proposto por este trabalho, foi preciso estabelecer alguns objetivos específicos que conduzissem às proposições necessárias ao processo de construção de taxonomias em portais corporativos.

O primeiro desses objetivos - caracterização dos modelos de representação do conhecimento - foi alcançado com o levantamento bibliográfico referente aos diferentes modelos. Ao se reconhecer que a classificação, o tesauro, a ontologia e a taxonomia apresentam diferenças quanto ao desenvolvimento e à aplicação, foi possível identificar a taxonomia como o modelo de representação que atende às necessidades do ambiente tanto de portais quanto do ambiente corporativo. Isso porque possibilita ao usuário fazer parte do processo de construção, apresentando uma estrutura flexível que atenda às suas necessidades informacionais.

Como segundo objetivo específico da pesquisa, buscou-se identificar as práticas de elaboração para construção de taxonomias disponíveis na literatura da área de Ciência da Informação. Foram identificados seis práticas de elaboração que apresentam etapas estabelecidas para este processo; cinco na língua inglesa e um na língua portuguesa. Com base nessa identificação, foi realizado um 'recorte' nas práticas de elaboração escolhidas, extraindo-se apenas as etapas de construção da taxonomia.

Quanto à caracterização das funcionalidades de um portal corporativo relacionadas à organização e representação da informação, pelo levantamento bibliográfico na área de portais corporativos, identificou-se a funcionalidade 'categorização'. A literatura da área de portais corporativos mostra que o conceito de categorização envolve tanto aspectos da construção de taxonomia (representação do conhecimento) como da aplicação da taxonomia na representação da informação e organização da informação recuperada. Sendo assim, categorização foi a variável central de análise desta pesquisa

No processo de análise das práticas de elaboração para a construção de taxonomias, levamos em consideração essa variável. Verificamos que, apesar de os portais corporativos indicarem as taxonomias como elementos centrais no processo de representação e organização da informação, a literatura da área não apresenta indicações metodológicas claras em relação ao seu desenvolvimento. Esse problema ocorre principalmente com a etapa da construção das relações semânticas entre os termos. Conforme observado, essa questão pode estar relacionada ao fato de as práticas de elaboração serem, na grande maioria das vezes, desenvolvidas por profissionais da área tecnológica, sem o auxílio de um profissional da informação, considerado mais qualificado para esse processo. A taxonomia constitui-se em um processo de classificação pois, segundo a FID/CR - Comitê Técnico de Pesquisa de Classificação (1973 apud CAMPOS, 2001), "classificação é qualquer método de reconhecimento de relações, genéricas ou outras, entre itens de informação, não importa o grau de hierarquia usada, nem se aqueles métodos são aplicados em conexão com sistemas tradicionais ou computadorizados de informação." Podemos constatar que a taxonomia trabalha os conceitos básicos da classificação aplicados em um ambiente diferenciado - a Internet. Desse modo, para a construção de taxonomias, é necessário resgatar conceitos relacionados a sistemas de classificação.

Como já mencionado, está implícito na literatura da área da taxonomia que a única organização possível é a hierárquica. Observamos, pela análise das necessidades dos portais corporativos e dos objetivos das taxonomias, que essa única organização seria insuficiente. Pela complexidade das relações estabelecidas em ambientes corporativos, é preciso ampliar a rede de relacionamentos conceituais. A informação não pode ser reduzida a relacionamentos genéricos e partitivos. Nas taxonomias corporativas, a informação precisa fazer sentido para os seus usuários, e esse propósito pode ser alcançado pelas relações seqüenciais, que não estão submetidas a uma relação de hierarquia, mas apresentam contigüidade espacial e temporal.

O exame das práticas de elaboração selecionadas na literatura possibilitou o estabelecimento de orientações para o processo de construção de taxonomias corporativas. Foram propostas e detalhadas as seguintes etapas:

1. Estabelecimento das categorias gerais.
2. Coleta dos termos.
3. Análise dos termos selecionados.
4. Controle da diversidade de significação.
5. Construção dos relacionamentos semânticos.

Verificamos que, além de apoiar-se na literatura da área de portais corporativos, o processo precisa estar respaldado com as teorias e métodos elaborados pelas áreas de Classificação, Indexação, Lingüística e Documentação. Essas teorias e métodos fornecem o devido aprofundamento teórico e metodológico necessário nas etapas da construção.

As limitações da pesquisa abrem perspectivas e precedentes para trabalhos futuros. Com o seu desenvolvimento, novas possibilidades de observação, entendimento e investigação surgem, trazendo novas questões.

A taxonomia, como modelo de organização da informação e representação do conhecimento da área da Ciência da Informação, é um tema relativamente recente, embora trabalhe com conceitos básicos da área. A literatura predominante sobre o tema ainda é desenvolvida por empresas de consultoria estadunidenses, com poucos avanços teóricos.

Assim, seria interessante que novas pesquisas buscassem aprofundar outras etapas do desenvolvimento de taxonomias não apresentadas neste trabalho. Outra

possibilidade seria a aplicação das proposições em um estudo de caso, buscando testá-las e aperfeiçoá-las.

Poderiam ser empreendidas pesquisas que extrapolassem o ambiente corporativo, direcionando a construção da taxonomia para outras áreas, como grupos de pesquisas em universidades. Essas novas alternativas proporcionam um redimensionamento da aplicação de taxonomias, emergindo possivelmente novos interesses de ordens práticas e teóricas.

REFERÊNCIAS

ABREU, A.F. ; ABREU, P.F. *Sistemas de informações gerenciais: uma abordagem orientada à gestão empresarial*. Florianópolis, 2002. Trabalho não publicado.

ADAMS, K.C. Immersed in structure: the meaning a function of taxonomies. *Internetworking*, USA, n.3.2, ago. 2000. Disponível em: http://www.internettg.org/newsletter/aug00/article_structure.html. Acesso em: 10 dez. 2005.

ALMEIDA, M.B. *Taxonomias, tesaurus ou ontologias*. 2005. Disponível em: http://www.eci.ufmg.br/mba/p1_2_0.html. Acesso em: 05 fev. 2006.

_____; BAX, M.P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ci. Inf.*, Brasília, v.32, n.3, p.7-20, set./dez. 2003.

ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempos e espaços digitais. *Enc. Bibli*, Florianópolis, n.15, 2003. Disponível em: http://www.encontros-bibli.ufsc.br/Edicao15/alvarenga_representacao.pdf. Acesso em: 28 jan. 2006.

ARAÚJO, C.A.A. Fundamentos teóricos da classificação. *Enc. Bibli*, Florianópolis, n.22, 2. sem. 2006. Disponível em: <http://www.encontros-bibli.ufsc.br/regular.html>. Acesso em: 25 out. 2006.

ARCE, M.V.S.; PÉREZ, T.S. Las comunidades virtuales y los portales como escenarios de gestión documental y difusión de información. *Anales de Documentación*, n.4, 2001. Disponível em: <http://www.um.es/fccd/anales>. Acesso em: 02 fev. 2006.

BALDINGER, K. *Teoria semántica: hacia una semántica moderna*. Madrid: Alcalá, 1970.

BARBIERI, C. *BI – Business Intelligence: modelagem e tecnologia*. Rio de Janeiro: Axcel, 2001.

BARDIN, L. *Análise de Conteúdo*. Lisboa: Edições 70, 2004.

BENETT, G. *Intranets: como implantar com sucesso na sua empresa*. Rio de Janeiro: Campus, 1997.

BLACKBURN, B. Taxonomy design types. *AJIM E-doc Magazine*, Maryland, USA. v.20, n.3, p.14-16, maio/jun. 2006.

CAFÉ, Ligia Maria Arruda. *Análise da informação*. Florianópolis, 2006. Trabalho não publicado.

CAMPOS, M.L.A. *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói: EdUFF, 2001.

_____. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ci. Inf.*, Brasília, v.33, n.1, p. 22-32, jan./abr. 2004.

_____; GOMES, H.E. Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador. *Perspect. Ci. Inf.*, Belo Horizonte, v.11, n.3, p.348-359, set./dez. 2006.

CARDOSO, C.A. *Porquê um portal organizacional (EIP)*. 2002. Disponível em: <http://www.kmol.online.pt/artigos/200201/car01_p.html. Acesso em: 02 dez. 2005.

CARVALHO, R.B.; FERREIRA, M.A.T. Análise das funcionalidades de intranets e portais: pesquisa exploratória em médias e grandes organizações brasileiras. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 6., 2005, Florianópolis. *Anais...* Florianópolis: PGCIN/UFSC, 2005. 1 CD-ROM.

CASANOVA, M. A. *Tecnologias de banco de dados para a web semântica: módulo 4a - ontologias – requisitos*. Rio de Janeiro: PUC, 2003. Disponível em: <<http://www.inf.puc-rio.br/~casanova/Topicos-WebBD-Notas/modulo4a-ontologias-requisitos-2.pdf> >. Acesso em 01 dez. 2005.

CAVALCANTI, C.R. *Indexação e tesauro: metodologia e técnicas*. Brasília: Associação dos bibliotecários do Distrito Federal, 1978.

CAVALCANTI, M.; GOMES, E.; PEREIRA, A. *Gestão de empresas na sociedade do conhecimento: um roteiro para a ação*. Rio de Janeiro: Campus, 2001.

CBO – Classificação Brasileira de Ocupações. 2002. Disponível em: <http://www.mteco.gov.br>. Acesso em: 11 jun. 2007.

CINTRA, A.M.M.; et al. *Para entender as linguagens documentárias*. São Paulo: Polis, 2002.

CLASSIFICATION RESEARCH GROUP. The need for a faceted classification as the basis of all methods of information retrieval. *Library Association Record*, London, v.57, n.7, p.262-268, 1955.

CONWAY, Susan; SLIGAR, Char. Building taxonomies. In.: _____. *Unlocking knowledge assets*. Redmont: Microsoft Press, 2002. Cap. 6. Disponível em: <<http://www.microsoft.com/mspress/books/sampchap/5516a.asp#100>. Acesso em: 04 set. 2006,

CHOO, C.W. *A organização do conhecimento: como as organizações usam a informação para criar significados, construir conhecimento e tomar decisões*. São Paulo: SENAC, 2003.

_____. Working with knowledge: how information professional help organisations manage what they know. *Library Management*, v.21, n.8, p.395-403, 2000.

DAVENPORT, T.H.; PRUSAK, L. *Conhecimento empresarial: como as organizações gerenciam o seu capital intelectual*. Rio de Janeiro: Campus, 1999.

DAHLBERG, I. Fundamentos teórico-conceituais da classificação. *Revista de Biblioteconomia de Brasília*, v.6, n.1, p.9-21, jan./jun. 1978a.

_____. Teoria do conceito. *Ci. Inf.*, Brasília, v. 7, n. 2, p.101-107, 1978b.

DELPHI GROUP. *Taxonomy and content classification*. 2002. Disponível em: www.delphigroup.com/research/whitepapers/WP_2002_TAXONOMY.pdf. Acesso em: 15 mar. 2006.

DESLANTES, S.F.; et al. *Pesquisa social: teoria, método e criatividade*. Petrópolis: Vozes, 1994.

DETLOR, Brian. The corporate portal as information infrastructure: towards a framework for portal design. *International Journal of Information Management*, v.20, n.2, apr. 2000.

DIAS, C.A. Portal corporativo: conceitos e características. *Ci. Inf.*, Brasília, v. 30, n. 1, p.50-60, jan./abr. 2001.

_____. *Usabilidade na Web: criando portais mais acessíveis*. Rio de Janeiro: Alta Books, 2003.

DODEBEI, V.L.D. *Tesouro: linguagem de representação da memória*. Niterói: Intertexto; Rio de Janeiro: Interciências, 2002.

DUDZIAK, E.A. ; VILLELA, M.C.O.; GABRIEL, M.A. *Gestão do conhecimento em bibliotecas universitárias*. Fortaleza: SNBU, 2002. Disponível em: <http://www.sibi.ufrj.br/snbu/snbu2002/oralpdf/91.a.pdf> Acesso em: 09 fev. 2004.

DUTRA, Jayne; BUSCH, Joseph. *Enabling knowledge discovery: taxonomy development for NASA*. 2003. Disponível em: <http://www.xml.gov/documents/completed/nasa/index.html>>. Acesso em: 04 out. 2006.

EARLEY AND ASSOCIATES. *Developing enterprise taxonomies*. 2003. Disponível em: http://www.earley.com/Earley_Report/ER_Taxonomy.htm. Acesso em: 04 out. 2006.

EDOLS, L. Taxonomies are what? *Free Pint*, Reino Unido, n.97, out. 2001. Disponível em: <http://www.freepint.com/issues/041001.htm#feature>>. Acesso em: 10 nov. 2005.

FEITOSA, A.L.G. *A integração entre sistemas legislativos, terminologia e web semântica na organização e representação da informação legislativa*. 2005. 0 f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2005.

GARSHOL, L.M. Metadata?Thesauri?Taxonomies?Topic Maps! *Journal of Information Science*, v.30, n.4, 2004. Disponível em: <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>. Acesso em: 10 fev. 2006.

GIL, A.C. *Como elaborar projetos de pesquisa*. São Paulo: Atlas, 2002.

GILCHRIST, A. Corporate taxonomies: report on a survey of current practice. *Online Information Review*, v.25, n.2, p. 94-102, 2001.

GOMES, H.E. A organização do conhecimento diante das novas tecnologias da informação. In: SIMPÓSIO: ESTADO ATUAL E PERSPECTIVAS DA CDU. 1996. *Simpósio...* Brasília : IBICT, 1996, p. 54-57.

GRUBER, T. *What is an ontology?* [s.l.:s.n.], 1996. disponível em:<<http://www-ksl.stanford.edu/kst/what-is-na-ontology.html>>. Acesso em: 03 dez. 2005.

HARVARD BUSINESS REVIEW. *Gestão do conhecimento – on knowledge management*. Rio de Janeiro: Editora Campus, 2000.

HOLGATE, L. Creating and using taxonomies to enhance enterprise search. *Information Today*, Medford, N.J., v.7, n.21, jul./ago. 2004.

INTERNACIONAL STANDARDS ORGANIZATION. *ISO 2788: Guidelines for the establishment and development of monolingual thesauri*. Geneve, 1974.

KREMER, S.; KOLBE, L.M.; BRENNER, W. Towards a procedure model in terminology management. *Journal of documentation*, Bradford:UK, v.61, n.2, p.281-295, 2005.

LANCASTER, F.W. *Construção e uso de tesauro: curso condensado*. Tradução de César Almeida de Meneses Silva. Brasília: IBICT, 1987.

LANGRIDGE, D. *Classificação: abordagem para estudantes de biblioteconomia*. Rio de Janeiro: Interciência, 1977.

LEMOS, C. Inovação na era do conhecimento. In.: LASTRES, M.M. H.; ALBAGLI, S.. *Informação e globalização na era do conhecimento*. Rio de Janeiro: Campus, 1999. Cap. 5. p.122-144.

LIMA, V.M.A. *Da classificação do conhecimento científico aos sistemas de recuperação de informação: enunciação de codificação e enunciação de decodificação da informação documentária*. 2004. 117f. Tese (Doutorado em Ciência da Comunicação) – Universidade de São Paulo, São Paulo, 2004.

LOJKINE, J. *A revolução informacional*. São Paulo: Cortez, 2002.

LYONS, J. Semântica. In.: _____. *Linguagens e lingüística: uma introdução*. Rio de Janeiro: Guanabara Koogem, 1987. Cap. 5. p. 133-146.

MARCONDES, C.H. Representação e economia da informação. *Ci. Inf.*, Brasília, v.30, n.1, p.61-70, jan./abr. 2001.

MARTINEZ, A.; et al. Las categorías o facetas fundamentales: una metodología para el diseño de taxonomías corporativas de sitios Web argentinos. *Ci. Inf.*, Brasília, v.33, n.2, p.106-111, maio/ago. 2004.

MONTAGUE INSTITUTE. *Managing taxonomies strategically*. 2001. Disponível em: <http://www.montague.com/abstracts/taxonomy3.html>>. Acesso em: 26 out. 2006.

MORANTE, Marcia. *Creating useful taxonomies: metadata, taxonomies and controlled vocabularies*. 2003. Disponível em: <http://www.kcurve.com/Metadata_Taxonomy%20Development_SLA_060804.ppt>. Acesso em: 03 nov. 2006.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A.P. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. *DataGramaZero*, v.5, n.6, dez. 2004. Disponível em: www.dgz.org.br/dez04/Art_01.htm. Acesso em 18 mar. 2006.

NATIONAL INFORMATION STANDARDS ORGANIZATION (U.S.). *Guidelines for the construction, format and management of monolingual thesauri* – developed by the National Information Standards Institute – ANSI/NISO Z39.19 – 2003 (revision of Z39.19-1980). Bethesda (USA): Niso Press, 2003

NISEMBAUM, H. Gestão do conhecimento: enriquecendo o capital humano. In: BOOG, Gustavo; BOOG, Magdalena (Coord.). *Manual de gestão de pessoas e equipes*. São Paulo: Gente, 2002. cap. 18, p. 275-286.

NONAKA, I. A empresa criadora do conhecimento. In: Harvard Business Review. *Gestão do conhecimento*. Rio de Janeiro: Campus, 2000.

_____; TAKEUCHI, H. *Criação de conhecimento na empresa: como as empresas japonesas geram a dinâmica da inovação*. Rio de Janeiro: Campus, 1997.

NOY, F. N.; GUINNESS, D.L. *Ontology development 101: a guide to create your first ontology*. 2001. Disponível em: <<http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>>. Acesso em: 03 dez. 2005.

PARREIRAS, F. *Ontologias em portais corporativos: fatores determinantes de sucesso*. 2004. Disponível em: <http://www.fernando.parreiras.nom.br/content/view/38/0/>. Acesso em: 01 dez. 2005.

PIEPADE, M.A.R. *Introdução à teoria da classificação*. Rio de Janeiro: Interciência, 1977.

PLOSKER, G. Taxonomies: facts and opportunities for information professionals. *Online. ABI/Inform Global*, v.1, n.29, jan./fev. 2005. p. 58-60.

RAMALHO, R.A.S.; VIDOTTI, S.A.B.G.; FUJITA, M.S.L. Web semântica: aspectos interdisciplinares para a organização e recuperação de informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 6., 2005, Florianópolis. *Anais...* Florianópolis: PGCIN/UFSC, 2005. 1 CD-ROM.

RANGANATHAN, S.R. Facet analysis: fundamental categories. 1962. In.: CHAN, L.M. *Theory of subject analysis*. Colorado: Libraries Unlimited, 1985.

REZENDE, D. A.; ABREU, A.F. *Tecnologia da informação aplicada a sistemas de informação empresariais*: o papel estratégico da informação e dos sistemas de informação nas empresas. São Paulo: Atlas, 2003.

SENGE, P. *A quinta disciplina: arte e prática da organização que aprende*. São Paulo: Best Seller, 1998.

SILVA, O.P.; GANIM, F. *Manual da CDU*. Brasília: Briquet de Lemos, 1994.

STEWART, T. *Capital intelectual: a nova vantagem competitiva das empresas*. Rio de Janeiro: Campus, 1998.

TEIXEIRA FILHO, J. *Gerenciando conhecimento*: como a empresa pode usar a memória organizacional e a inteligência competitiva no desenvolvimento de negócios. Rio de Janeiro: Editora SENAC, 2001.

TERRA, J.C.C.; et al. *Taxonomia: elemento fundamental para a gestão do conhecimento*. 2005. Disponível em : <http://www.terraforum.com.br>. Acesso em: 12 nov. 2005.

_____; BAX, M.P. Portais Corporativos: instrumento de gestão de informação e de conhecimento. In.: PAIM, I. (Org.) *A gestão da informação e do conhecimento*. Belo Horizonte: ECI/UFMG, 2003. p.33-53.

_____; GORDON, C. *Portais corporativos: a revolução na gestão do conhecimento*. São Paulo: Elsevier, 2002.

TURBAN, E.; RAINER JUNIOR, R.K.; POTTER, R.E. *Administração de tecnologia da informação: teoria e prática*. Rio de Janeiro: Elsevier, 2003.

VAN REES, R. *Clarity in the usage of the terms ontology, taxonomy and classification*. 2003. Disponível em: <http://vanrees.org/research/papers/Cib78ConferencePaper2003>. Acesso em: 20 nov. 2005.

VICKERY, B.C. *Faceted classification: a guide to the construction and use of special schemes*. Londres: Aslib, 1960.

WELTY, C.A. *An integrated representation for software development and discovery*. 1996. Disponível em: <http://www.cs.vassar.edu/faculty/welty/papers/phd/HTML/dissertation-1.html>. Acesso em: 01 dez. 2005.

WOODS, E. The corporate taxonomy: creating a new order. *KMWorld*, USA, v.13, n.7, jul., 2004. Disponível em: <http://www.kmworld.com/Articles/ReadArticle.aspx?ArticleID=9566>. Acesso em: 15 mar. 2006.

WORLD-WIDE WEB CONSORTIUM – W3C. Disponível em: <http://www.w3.org>. Acesso em: 20 fev. 2006.

YAMAOKA, E.J. *Taxonomia e metadados como elementos estruturantes da organização e representação do conhecimento numa empresa*. 2005. 151f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2005.