

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO

MÁRCIA MARTINS DE ARAÚJO ALTOUNIAN

**MODELO CONCEITUAL PARA UMA ONTOLOGIA DE LICITAÇÕES E
CONTRATOS ADMINISTRATIVOS: PROPOSTA DE METODOLOGIA PARA O
TRIBUNAL DE CONTAS DA UNIÃO**

Brasília

2013

MÁRCIA MARTINS DE ARAÚJO ALTOUNIAN

**MODELO CONCEITUAL PARA UMA ONTOLOGIA DE LICITAÇÕES E
CONTRATOS ADMINISTRATIVOS: PROPOSTA DE METODOLOGIA PARA O
TRIBUNAL DE CONTAS DA UNIÃO**

Monografia apresentada ao programa de Especialização do Núcleo de Informação Tecnológica e Gerencial – NITEG, no curso de Especialização em Arquitetura e Organização da Informação da Escola de Ciência da Informação, da Universidade Federal de Minas Gerais, como requisito para a obtenção do certificado de Especialista em Arquitetura e Organização da Informação.

Orientador: Prof.^a Dr.^a Maria Aparecida Moura

BRASÍLIA

2013

ALTOUNIAN, Márcia Martins de Araújo

Modelo conceitual para uma ontologia de Licitações e Contratos Administrativos: proposta de metodologia para o Tribunal de Contas da União/Márcia Martins de Araújo Altounian. Brasília, 2013.

73 p.

1. Tesauro 2. Ontologia 3. Modelo conceitual 4. Ciência da Informação 5. Teoria do Conceito. II. Título.

Para Cláudio, Rodrigo e Pedro Felipe.

AGRADECIMENTOS

Agradeço a Júnia Beatriz de Oliveira Souza e Clemens Soares dos Santos, do Centro de Documentação do Tribunal de Contas da União, pelo incentivo e apoio;

A Flávia Lacerda, pela viabilização da minha participação no curso e pelas orientações e sugestões;

A Câmara dos Deputados, especialmente ao Diretor do Centro de Documentação e Informação, Adolfo Furtado, pela oferta de vagas do curso ao Tribunal de Contas da União;

Aos professores da Escola de Ciência da Informação, da Universidade Federal de Minas Gerais;

Aos colegas da Câmara dos Deputados, pela acolhida.

RESUMO

No atual cenário de crescimento da demanda por sistemas de informação que permitam um melhor compartilhamento e recuperação de informação e conhecimento, as ontologias vêm se destacando como ferramenta de grande potencial para o alcance desse objetivo. Originado na Filosofia, o termo foi adotado pela Ciência da Computação (CC) e tem sido alvo de estudos na área de Ciência da Informação (CI). O trabalho trata o tema sob a perspectiva da Ciência da Informação, a partir da identificação dos fundamentos teóricos que possam contribuir na construção de ontologias, especificamente na fase de modelagem conceitual, e aponta qual o mais adequado para esse fim. A partir dessa constatação é proposta uma diretriz metodológica a ser utilizada na construção de uma ontologia no domínio de Licitações e Contratos Administrativos para o Tribunal de Contas da União.

Palavras-chave: Tesouro, Ontologia, Modelo conceitual, Ciência da Informação, Teoria do Conceito.

LISTA DE FIGURAS

Figura 1	Atividades de inteligência do controle_____	19
Figura 2	Triângulo do conceito_____	45
Figura 3	Comparação de conceitos do ponto-de-vista da lógica formal____	46
Figura 4	Categorias e subcategorias de Dahlberg_____	47

LISTA DE QUADROS

Quadro 1	Quadro comparativo - metodologias de modelagem conceitual de ontologias na Ciência da Computação e na Ciência da Informação_____	55
----------	--	----

LISTA DE ABREVIATURAS E SIGLAS

ANSI/NISO - American National Standard / National Information Standards Organization

CC – Ciência da Computação

CI - Ciência da Informação

CRG - Classification Research Group

DW – Datawarehouse

ER – Modelo Entidade Relacionamento

IA – Inteligência Artificial

IEEE - Institute of Electrical and Electronics Engineer

ISO – International Organization for Standardization

LD – Linguagem documentária

PMEST – Propriedade, matéria, energia, espaço e tempo

SI – Sistema de informação

SOC – Sistema de Organização do Conhecimento

TCU – Tribunal de Contas da União

TECON – Tesouro do Tribunal de Contas da União

UML - Unified Modeling Language

W3C - World-Wide Web Consortium

SUMÁRIO

1.	Introdução_____	12
1.1	Problema_____	14
1.2	Objetivos_____	15
1.2.1	Objetivo geral_____	15
1.2.2	Objetivos específicos_____	15
1.3	Justificativa_____	16
2	Tribunal de Contas da União e controle externo _____	17
2.1	Fiscalização de licitações e contratos administrativos_____	17
2.1.1	Repositório de conhecimento de controle externo_____	18
3	Metodologia_____	21
4	Referencial teórico_____	22
4.1	Representação do conhecimento e linguagens documentárias_____	22
4.2	Classificação_____	23
4.3	Taxonomia_____	26
4.4	Tesouro_____	27
4.5	Ontologia_____	29
4.6	Modelagem conceitual_____	34
4.6.1	Modelagem conceitual em ontologias_____	37
4.7	Fundamentos da Ciência da Informação aplicados à construção de ontologias_____	39
4.8	Teoria do conceito_____	44

5	Proposta de metodologia de modelagem conceitual para uma ontologia no domínio licitações e contratos administrativos _____	52
5.1	Metodologia proposta _____	53
5.1.1	Definição do escopo da ontologia _____	54
5.1.2	Aquisição de conhecimento _____	54
5.1.3	Identificação de termos _____	54
5.1.4	Definição dos conceitos _____	55
5.1.5	Agrupamento dos termos em categorias _____	55
5.1.6	Construção da estrutura taxonômica _____	56
5.1.7	Definição das relações _____	56
5.1.8	Definição das instâncias _____	57
6	Resultados _____	59
7	Conclusão _____	61
8	Referências _____	63

INTRODUÇÃO

O atual cenário tecnológico demanda o desenvolvimento e aprimoramento de sistemas e metodologias que trabalhem com a recuperação e o intercâmbio de informação e conhecimento. Nesse ambiente, a clareza conceitual representa elemento basilar para o alcance desse desafio contemporâneo.

O aumento exponencial do volume de informações disponibilizadas em meio digital torna urgente um olhar sobre a dimensão dos modelos convencionais e dos modelos potenciais de representação do conhecimento.

Novos tipos de recursos informacionais têm se apresentado à sociedade, como: livros eletrônicos (*e-books*), documentos digitalizados em texto integral, *sites web*, monografias, periódicos, teses e dissertações em formato digital, listas de discussão, blogs, *wikis* e redes sociais. Diante dessa recente perspectiva, as ontologias têm-se destacado como uma nova possibilidade para a representação do conhecimento.

Originado na Filosofia, o termo “ontologia” foi adotado pela Ciência da Computação (CC) e tem sido alvo de estudos na área de Ciência da Informação (CI). A ontologia pretende viabilizar a definição de um conjunto de conceitos, relações e propriedades, para que um sistema possa explicitar o contexto e o significado dos termos adotados e realizar inferências. Essa tarefa depende do uso de uma linguagem formal, livre de ambiguidades, as quais o computador não pode resolver como fazem as pessoas em suas interações através da linguagem natural. Tais inferências são primárias se comparadas com as possibilidades humanas, mas significam avanço para a automatização de ações no ambiente digital.

Por serem relativamente recentes, as ontologias ainda são alvo de dissenso dentro da CC, onde se observa a carência de metodologias sólidas para sua construção, devido principalmente às restrições de representatividade semântica. Por outro lado, a CI com sua tradição na construção de sistemas de organização do conhecimento, como classificações, taxonomias e tesauros possui rico arcabouço teórico sobre a construção de linguagens documentárias.

A intenção do presente trabalho é tratar o tema sob a perspectiva da Ciência da Informação e propor o aproveitamento de seu arcabouço teórico e metodológico como subsídio na construção de uma ontologia.

Especificamente, tratar-se-á do assunto propondo-se a adoção das bases metodológicas utilizadas na construção de tesauros na modelagem conceitual no domínio de Licitações e Contratos Administrativos, apontando-se, dentre as metodologias da CI, a mais adequada para esse fim.

A partir dessa investigação, pretende-se propor uma metodologia de construção de ontologias para o Tribunal de Contas da União (TCU), que contribua com o aprimoramento do controle externo e com o compartilhamento de conhecimento dentro e fora da instituição e que possa, ainda, ser aplicada em outros domínios de interesse da Corte de Contas.

1.1 PROBLEMA

O grande interesse em ontologias, atualmente, dá-se pela necessidade de atender a crescente busca pelo maior troca e reutilização de informações entre os sistemas e as pessoas. Para o compartilhamento de informações é fundamental a padronização de conceitos e da terminologia dentro de um mesmo domínio. A ausência de uma padronização dificulta o intercâmbio de informações entre diferentes sistemas, bem como a interoperabilidade e o reuso. Torna-se, portanto, indispensável a codificação do conhecimento para viabilizar a acessibilidade por parte das pessoas que dele precisam. Assim também ocorre dentro das organizações, que precisam de uma codificação ou representação formal e explícita do conhecimento a fim de tratá-lo adequadamente com vistas ao compartilhamento e reuso.

No TCU, a necessidade de se construir uma padronização de definições e significados, que aqui denominaremos de modelo conceitual, nasceu da demanda de se agregar semântica a um sistema de inteligência corporativo, especificamente no domínio de Licitações e Contratos Administrativos.

Uma ontologia é composta de algumas fases, sendo uma delas a modelagem conceitual, que consiste em criar um modelo representativo da realidade de um determinado domínio, por meio de conceitos e das relações entre eles. Ela também é composta de um vocabulário lógico – linguagem formal e compreensível pela máquina - e um vocabulário terminológico – linguagem natural e compreensível pelo homem.

Para uma modelagem conceitual consistente é necessário que tal atividade seja conduzida com métodos bem estruturados a fim de garantir que a ontologia cumpra seu papel de reutilização e compartilhamento de informação e conhecimento.

Considerando que a CI possui um arcabouço teórico bastante consistente, no que diz respeito às metodologias e técnicas para a representação da informação e do conhecimento, tais práticas podem se mostrar pertinentes na fase de modelagem conceitual de um sistema de organização do conhecimento em meio digital.

Pelo exposto, a questão que se coloca no presente estudo é: **As metodologias de modelagem conceitual da CI podem contribuir para estruturar o vocabulário terminológico de uma ontologia na área de Licitações e Contratos Administrativos?**

1.2 OBJETIVOS

1.2.1 Objetivo geral:

Propor uma metodologia de modelagem conceitual para estruturar o vocabulário terminológico de uma ontologia aplicada ao domínio de gestão pública na área de Licitações e Contratos Administrativos.

1.2.2 Objetivos específicos:

- Identificar e sistematizar os fundamentos teóricos da área de Organização da Informação e do Conhecimento que possam contribuir para a construção de ontologias aplicáveis a diferentes domínios de conhecimento;
- Identificar, selecionar e adequar metodologias de modelagem conceitual fundamentadas na Ciência da Informação adequadas para estruturar o vocabulário terminológico de uma ontologia aplicada ao domínio de gestão pública na área de Licitações e Contratos Administrativos.

1.3 JUSTIFICATIVA

O alinhamento entre a Tecnologia da Informação e as áreas de negócio nas organizações tem sido uma premissa atual consistente para o alcance dos objetivos institucionais.

Na Arquitetura Corporativa, observa-se a preocupação crescente com um dos seus pilares, a Arquitetura da Informação, que visa prover uma visão integrada da informação que trafega entre os processos organizacionais. As ontologias se inserem nesse cenário, assim como taxonomias, vocabulários controlados e metadados, posicionando-se como Estruturas de Representação. Todas essas ferramentas são modelos de representação do conhecimento que possuem a função de organizar, facilitar a navegação, indexar, representar e atribuir significado às informações disponíveis de modo a facilitar sua recuperação.

A construção de relacionamentos semânticos entre conceitos representados em ontologias de domínio e elementos dos processos de negócio, dos modelos lógicos de dados e do esquema físico de bases de dados promove uma melhor arquitetura de informação dentro de uma organização. Uma ontologia permite gerir o conhecimento da organização, trocar informação com outros sistemas de informação, recuperar e reutilizar conhecimento de um domínio e controlar de forma efetiva grande quantidade de informação.

Diante da potencialidade dessa ferramenta, a possibilidade de tratar do tema mostrou-se pertinente ao propósito do trabalho de conclusão do curso de Especialização em Arquitetura e Organização da Informação (AOI), além de guardar proximidade com a área de atuação profissional da aluna, a saber, linguagens documentárias. A oportunidade de aplicar os conhecimentos de construção de sistemas de organização do conhecimento (SOC) convencionais em ambientes digitais certamente foi decisiva para a escolha deste estudo.

Por fim, a viabilidade da parceria do Centro de Documentação (CEDOC) com a Diretoria de Informações Estratégicas (DGI) do TCU no desenvolvimento de um modelo conceitual para a integração de uma ontologia no domínio de Licitações e Contratos Administrativos, com o intuito de contribuir para a melhoria do desempenho de um sistema corporativo, ratificou a opção pelo tema ora tratado.

2. TRIBUNAL DE CONTAS DA UNIÃO E O CONTROLE EXTERNO

O TCU é órgão auxiliar do Congresso Nacional no exercício do Controle Externo e dentre suas competências, elencadas na Constituição Federal de 1988, art. 71 e regulamentadas pela Lei nº 8.443/1992 (Lei Orgânica), Lei nº 8.666/1993 (Lei de Licitações e Contratos Administrativos) e Res. TCU nº 246/2011 (Regimento Interno), incluem-se o julgamento das contas dos administradores e demais responsáveis por dinheiros, bens e valores públicos da administração direta e indireta e a fiscalização dos atos de que resulte receita ou despesa praticados pelos responsáveis sujeitos à sua jurisdição.

O controle é algo inerente à atividade de todas as pessoas e tem por significado, segundo o Dicionário Aurélio, “a fiscalização exercida sobre a atividade de pessoas, órgãos, departamentos ou sobre produtos, etc, para que tais atividades ou produtos não se desviem das normas preestabelecidas”.

Na Administração Pública existem categorias de controle, sendo uma delas o Controle Externo, que é função do Poder Legislativo, de competência do Congresso Nacional no âmbito federal, das Assembleias Legislativas nos Estados, da Câmara Legislativa no Distrito Federal e das Câmaras Municipais nos Municípios com o auxílio dos respectivos Tribunais de Contas. Trata-se de função fiscalizadora exercida pelo povo, por meio de seus representantes, sobre a administração financeira e orçamentária. É um controle de natureza política, mas sujeito à prévia apreciação técnico-administrativa do Tribunal de Contas competente.

Além disso, caracteriza-se por ser o controle onde há fiscalização de um Poder sobre os atos administrativos de outro Poder.

A fiscalização financeira e orçamentária é conferida em termos amplos ao Congresso Nacional, mas se refere fundamentalmente à prestação de contas de todo aquele que administra bens, valores ou dinheiros públicos. É decorrência natural da administração, como atividade exercida em relação a interesses alheios. Não é, pois, a natureza do órgão ou da pessoa que o obriga a prestar contas; é a origem pública do bem administrado ou do dinheiro gerido que acarreta para o gestor o dever de comprovar o seu zelo e bom emprego. Toda a administração pública - já o dissemos - fica sujeita a fiscalização hierárquica, mas certamente por sua repercussão imediata no erário, a administração financeira e orçamentária submete-se a maiores rigores de acompanhamento, tendo a Constituição da República determinado o controle interno pelo Executivo e o controle externo pelo Congresso Nacional auxiliado pelo Tribunal de Contas da União (MEIRELLES, 1990, p. 601).

2.1 Fiscalização de Licitações e Contratos Administrativos

Após a promulgação da Constituição Federal de 1988, diversas leis trouxeram novas referências e marcos para a atuação dos Tribunais de Contas.

No tocante às licitações e contratos, podem-se destacar a Lei nº 8.666, de 1993, que institui normas para licitações e contratos administrativos; a Lei nº 10.520, de 2002, que cria a modalidade de licitação denominada pregão; e a Lei 12.462, de 2011, que dispõe sobre o Regime Diferenciado de Contratações Públicas – RDC.

A licitação é o procedimento para a contratação na esfera da Administração Pública que objetiva a seleção da proposta mais vantajosa e a promoção do desenvolvimento nacional sustentável, apenas sendo dispensável ou inexigível nos casos expressamente previstos em lei. Além de um dever, a licitação é uma obrigação do gestor público, a qual impõe um padrão de comportamento administrativo para proporcionar oportunidades iguais a todos que desejem contratar com o Poder Público.

O art. 113 da Lei nº 8.666/1993 concede competência aos tribunais de contas para fazer o controle das despesas decorrentes dos contratos e demais instrumentos regidos por essa legislação. Ressalte-se que o *caput* deste artigo disciplina que cabe ao gestor comprovar que a sua atuação pautou-se pelos ditames legais. Com isso, pode-se dizer que, quando o órgão de controle encontrar indícios de ilegalidade, o ônus da prova competirá ao gestor.

2.1.1 Repositório de conhecimento de controle externo

Em face do arcabouço legal e da conseqüente responsabilidade em acompanhar a legalidade dos procedimentos licitatórios, o TCU tem se preocupado com a busca de novas técnicas para a seleção dos objetos auditados e com a produção e disseminação de conhecimentos na área. Nesse sentido, destaca-se a criação da Rede Interna de Informações, instituída pela Portaria TCU nº 205, de 2009, que se propõe a ser um canal de comunicação, sistematização e troca de informações entre todas as unidades técnicas que a integram. A referida rede constitui o suporte estratégico para a disseminação dos conceitos básicos necessários à utilização da atividade de inteligência voltada para o controle.

Com efeito, as ações de inteligência buscam trabalhar com o cruzamento e análise de grandes bases de dados disponíveis na administração pública federal, como o SIAFI (Sistema Integrado de Administração Financeira do Governo Federal), o SIASG (Sistema de Administração de Serviços Gerais) e o SIAPE (Sistema Integrado de Administração de Recursos Humanos), além de fontes de dados não estruturados, como o sistema de jurisprudência do TCU, notícias veiculadas pela imprensa e denúncias ou informações recebidas pela Ouvidoria do Tribunal. A partir de

parâmetros previamente estabelecidos (“tipologias de ilícitos administrativos”), é possível identificar ocorrência de irregularidades, em larga escala, em órgãos e entidades diversas.

Atividades de inteligência do controle

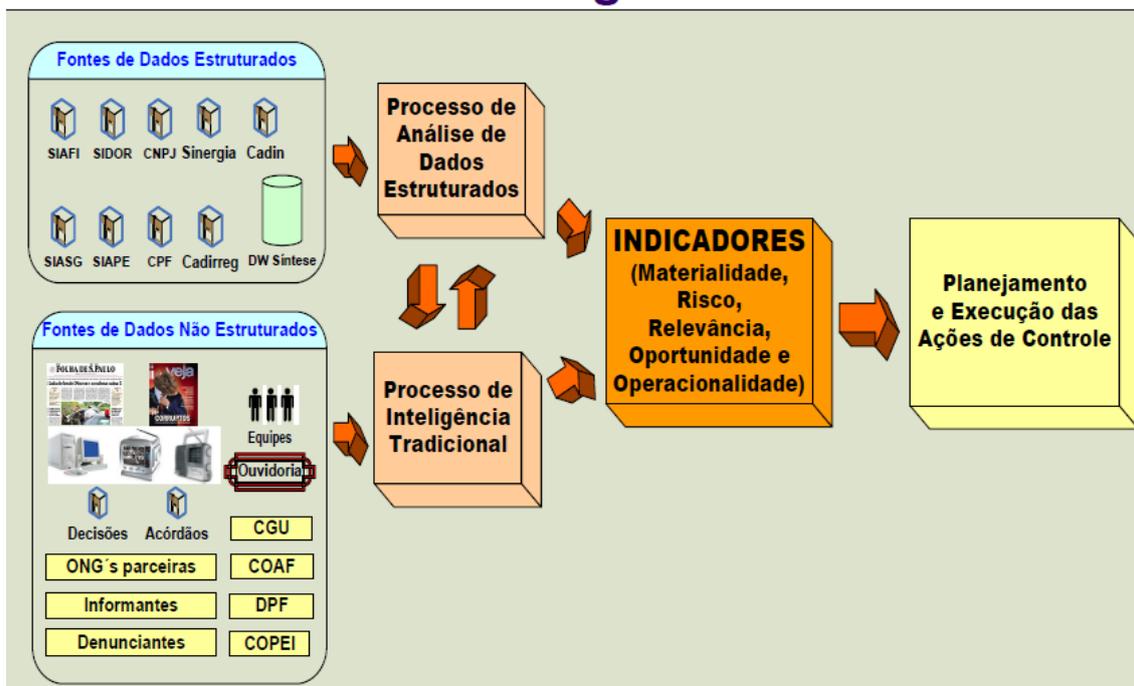


Figura 1: Atividades de inteligência do controle. Fonte: Código de ética dos servidores do TCU.

A simples existência e a disponibilidade de elevados volumes de dados de interesse para o controle não garantem a extração de indícios e provas relativas aos atos ilícitos cometidos contra o erário. O problema que se coloca é como lidar com a complexidade, diversidade e a gigantesca massa de informações disponíveis. A tarefa não é trivial, pois exige profissionais que agreguem perfil extremamente especializado que integre o conhecimento das regras de negócio e da estrutura das bases de dados consultadas.

A fim de lidar com esse desafio, buscou-se subsídios na área conhecida como Extração de Conhecimento ou KDD (*Knowledge-Discovery in Databases*), a chamada Mineração de Dados (*Data Mining*), que é o processo de identificação de padrões novos, potencialmente úteis e compreensíveis, em um conjunto de dados. Esses padrões podem se constituir, dentre outros, em regras de associação, ou sequências temporais que permitam revelar relacionamentos não aleatórios entre atributos de variáveis. O volume excessivo de dados justifica a mineração de dados para a descoberta de padrões que possam revelar informações úteis à tomada de decisão, pois podem produzir conhecimento valioso para o planejamento e a execução de fiscalizações.

Apesar dos algoritmos de mineração de dados atenderem a critérios de desempenho na criação de novos dados, muitas vezes não são suficientes para garantir seu uso prático. Conhecer o domínio é fundamental e as ontologias, que agregam conhecimento e semântica ao processo de mineração de dados, têm sido uma útil ferramenta para esse fim.

Uma das novas ações de inteligência do Tribunal para detectar indícios de fraudes e irregularidades em contratos firmados pela Administração Pública é o cruzamento de bases de dados governamentais, por meio do *datawarehouse* (DW) denominado Repositório de Conhecimento de Controle Externo. Atualmente, o desafio que se coloca ao DW utilizado pelo TCU é a combinação da mineração de dados com uma ontologia no domínio de Licitações e Contratos Administrativos para incrementar a *performance* do sistema.

Para superar esse desafio, é necessária a construção de uma ontologia, que apresenta diversas fases, dentre as quais se destaca a modelagem conceitual, objeto do presente trabalho. Nesta fase, o domínio deve ser retratado com a maior fidelidade possível, o que será viabilizado por meio de uma metodologia robusta. É neste ponto que a CI tem a oportunidade de contribuir com sua base teórica para a construção de um vocabulário terminológico que aprimore o desempenho do sistema e conseqüentemente, forneça informações relevantes para balizar as ações no TCU.

3. METODOLOGIA

A metodologia de pesquisa que fundamentou o presente trabalho é qualitativa de base teórica e teve como propósito consolidar um modelo conceitual aplicável à construção de uma ontologia.

A pesquisa que fundamentou o presente trabalho é aplicada, qualitativa descritiva e bibliográfica.

Aplicada quanto à natureza, pois pretende gerar conhecimentos para aplicação prática, com vistas a solucionar um problema específico.

Qualitativa porque realiza análise da literatura e propõe uma metodologia para construção de modelo conceitual em uma ontologia, sem, no entanto, 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”.

Descritiva, uma vez que o assunto já é conhecido e a contribuição é proporcionar uma nova visão sobre uma realidade já existente.

Por fim, bibliográfica, considerando que os subsídios teóricos utilizados foram coletados de fontes bibliográficas relacionadas à Ciência da Informação e Ciência da Computação.

4. REFERENCIAL TEÓRICO

4.1 Representação do conhecimento e linguagens documentárias

Representação pode ser definida como uma reprodução ou cópia de algum objeto ou fato. Quando se fala de representação da informação ou do conhecimento está se tratando de um processo complexo, uma vez que se tenta reproduzir em algum suporte, por meio de símbolos ou linguagens, conceitos e significados construídos ou percebidos pela mente humana (SILVA, 2008).

A organização da informação, segundo Café e Bräscher (2008), é um processo que envolve a descrição física e de conteúdo dos objetos informacionais, cujo produto descritivo é a representação da informação, entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico. Para as autoras, nesse contexto, o objeto é o registro da informação, que pertence ao mundo físico. Já a organização do conhecimento se aplica a unidades do pensamento (conceitos) e “[...] visa à construção de modelos de mundo que se constituem em abstrações da realidade” e apresenta como produto desse processo, a representação do conhecimento, que “[...] é fruto de um processo de análise de domínio e procura refletir uma visão consensual sobre a realidade que se pretende representar” (CAFÉ e BRÄSCHER, 2008, p. 6).

O conhecimento possui, em essência, uma natureza subjetiva, individual e intransferível, podendo ser somente elaborado por uma reflexão pessoal de alguém ao declarar que conhecimento é a “[...] certeza subjetiva ou objetivamente conclusiva da existência de um fato ou do estado de um caso, não sendo transferível e somente podendo ser adquirido por meio da reflexão” (DAHLBERG, 2006 *apud* RAMALHO, 2009, p. 3).

Entretanto, esse conhecimento subjetivo e individual poderá ser disseminado mediante formas de representação escrita ou falada, considerando-se nosso conhecimento prévio linguístico, possibilitando sua compreensão e a realização de comparações (FUJITA, 2008).

[...] de acordo com o paradigma de representação, os seres humanos são conhecedores e observadores da realidade externa. O processo de conhecimento consistiria na assimilação de coisas por meio de suas representações na mente/cérebro do sujeito cognoscente. Essas representações, uma vez processadas e codificadas no cérebro seriam comunicadas a outras mentes e/ou armazenadas e processadas por máquinas (computadores). [...] As informações seriam usadas para propósitos racionais específicos, mas, em nossos dias, tornar-se-ia cada vez mais aceitável a hipótese de que também as máquinas seriam capazes de alcançar níveis de “interpretação” e uso da informação além do simples processamento (CAPURRO, 1995 *apud* ALVARENGA, 2003, p. 21).

A grande dificuldade de se tratar o conhecimento, segundo Alvarenga (2003), é o fato de que a tentativa de classificar coisas, seres e textos sobre esses produzidos esbarra na constatação de que as coisas e os conhecimentos que lhes são correspondentes não se reduzem ao que deles pode-se ver explicitamente. Não é suficiente que se extraia somente características superficiais ou palavras deles constantes, é necessário um aprofundamento nas relações daquele objeto no contexto em que está inserido.

Os mecanismos de representação de conhecimento permitem que processos de formalização sobre os objetos e suas relações, em contextos predefinidos, possam ser facilmente representados (DAVIS, 1992 *apud* CAMPOS, 2004). Segundo Bräscher (2009, p. 4), essa representação é feita por meio de diferentes tipos de sistemas de organização do conhecimento (SOC), que representam determinado domínio por meio da sistematização dos conceitos e das relações semânticas que se estabelecem entre eles. O termo *knowledge organization systems* (KOS) foi proposto em 1998 pelo *Networked Knowledge Organization Systems Working Group* para englobar sistemas de classificação, tesouros, cabeçalhos de assunto, arquivos de autoridade, taxonomias, redes semânticas e ontologias.

Para Sales (2007), os modelos de representação do conhecimento, mais conhecidos como linguagens documentárias, são instrumentos de auxílio nas atividades de representação de conteúdos informacionais, visando o armazenamento de informações facilmente recuperáveis. Cada área do saber possui sua terminologia própria, que está restrita aos conhecedores de determinado conjunto de códigos, muitas vezes não sendo de conhecimento geral e que como linguagem de especialidade é concebida tanto pelos produtores das informações quanto por seus consumidores (usuários).

Sales e Café (2008) entendem as linguagens documentárias como exemplos de instrumentos que objetivam o controle dessa precisão terminológica no âmbito da representação do conhecimento e da recuperação da informação. A representação do conhecimento refere-se à ação de representar (no sentido de tornar presente) conteúdos de informações materializadas (registradas).

Cavalcanti (1978) entende linguagem documentária como um conjunto de regras, símbolos e termos pré-estabelecidos com o intuito de indicar assuntos de documentos.

Cintra (2002) faz a distinção entre linguagem natural (LN) e linguagem documentária (LD) ao afirmar que ambas são sistemas simbólicos instituídos que visam facilitar a comunicação, mas, a primeira é composta por palavras e seus diversos significados, e a segunda, é composta por palavras que assumem o papel de

termos, ou seja, restringem seus significados de acordo com um contexto ou área específica do conhecimento. A autora destaca ainda o aspecto social das linguagens documentárias, quando ressalta que além de simplesmente traduzirem o conteúdo de um documento, elas têm a função de possibilitar o compartilhamento social entre estoques de conhecimento e usuários de informação.

Qualquer LD é formada por termos, conceitos e características. De acordo com Sales (2007), os termos são considerados pela ISO 704 (2000) e pela ISO 1087-1 (2000) como sendo a unidade mínima da terminologia. Para Lara (2004) o termo é uma designação que corresponde a um conceito em uma linguagem de especialidade. Conceito, por sua vez, representa para Dahlberg (1978), a soma total dos enunciados verdadeiros sobre um objeto e Gomes (1990) o considera como um construto mental que representa um objeto individual material ou imaterial.

4.2 Classificação

Os seres humanos classificam a todo o momento as outras pessoas e os objetos do mundo [...] Todas as nossas ações no mundo são envolvidas por atos classificatórios, pois a ação de classificar é parte constitutiva das sociedades, na medida em que estão por toda parte, impregnam a vida social de forma onipresente (COSTA, 1997/1998 *apud* ARAÚJO, 2006).

Em termos históricos, é atribuído a Aristóteles o fato de ter sido o primeiro a propor a formulação de uma teoria da classificação. A sua contribuição inicial diz respeito à divisão dicotômica dos objetos em gênero e espécie. Trata-se de uma hierarquização conceitual que divide um tema geral em espécies a partir da aplicação de uma característica classificatória (ARAÚJO, 2006).

Barbosa (1969, p. 14) explica que para haver gêneros e espécies é imprescindível a existência de um princípio classificatório ou “característica de uma classificação”, que representa o “elemento que serve para reunir os grupos segundo as semelhanças que apresentam.” Essa é a segunda contribuição de Aristóteles para a teoria da classificação: a elaboração dos cinco predicados isto é, os cinco tipos de relações existentes num arranjo lógico:

- a) Gênero: classe ou grupo de seres ou objetos que possuem um determinado número de características em comum;
- b) Espécie: ser ou coisa que possui uma diferença específica que a distingue de seu gênero próximo; a espécie é obtida do gênero pelo acréscimo de uma diferença;
- c) Diferença: é a característica que serve para gerar uma espécie; cada acréscimo de diferença gera uma nova espécie;
- d) Propriedade: algo próprio de cada elemento de uma classe, mas que não é imprescindível à definição da classe;

e) Acidente: qualidade não obrigatória a todos os elementos de uma classe, isto é, que pode ou não estar presente em um conceito (BARBOSA, 1969, p. 14).

Conforme Pombo (1998, p. 2), há quatro grandes orientações: a ontológica (classificação dos seres), a gnosiológica (classificação das ciências), a biblioteconômica (classificação dos livros) e a informacional (classificação das informações).

Ao terceiro e quarto níveis (classificação dos livros e das informações) corresponde a constituição de uma ciência da classificação, isto é, de um novo domínio científico que tem por tarefa o estudo de todos os possíveis sistemas de classificação [...] (POMBO, 1998, p. 3).

De acordo com Silva (2008), os esquemas de classificação mais conhecidos até o início do século XX eram a *Library of Congress Classification*, a *Rider's International Classification*, a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU).

Na primeira metade do século XX, surge a Teoria da Classificação Facetada, desenvolvida por Shiyali Ramamrita Ranganathan. A inovação trazida foi que as teorias da classificação antes divididas conforme a finalidade (filosóficas e bibliográficas) e estas últimas, conforme a amplitude de aplicação (gerais ou enciclopédicas e especializadas) ou o tipo de característica (naturais e artificiais), após a teoria da classificação facetada também puderam ser divididas, de acordo com a forma de apresentação, em enumerativas (hierárquicas) e analítico-sintéticas¹ (ARAÚJO, 2006).

Segundo Barbosa (1969, p. 16), o diferencial do sistema classificatório proposto por Ranganathan é a utilização de uma estrutura dinâmica, multidimensional, com a introdução do termo faceta “que ficou sendo, nos modernos estudos sobre teoria da classificação, o substituto de característica”.

Baseando-se nas dez categorias do Ser, definidas por Aristóteles, [...] Ranganathan percebeu que todas elas poderiam ser utilizadas como características classificatórias, diferentemente do que ocorria até então nos sistemas hierárquicos e monolíticos, onde apenas uma categoria podia ser utilizada por vez como característica de classificação (ARAÚJO, 2006).

Partindo desse conceito de facetas, Ranganathan idealizou uma estrutura bem mais elástica do que as já existentes, que permitisse a síntese, o agrupamento de

¹Isto implica em que a classificação de cada documento exige uma análise de seu título ou um enunciado descritivo de seu conteúdo, nos termos dos conceitos que formam os componentes da ciência à qual esse documento pertence. Após a análise e a orientação desses elementos nas facetas, é possível sintetizá-las em expressões combinatórias que formam a classe, construída analítico-sinteticamente” (DAHLBERG, 1978).

vários componentes para a especificação do assunto. Para isso, adotou o uso de dois pontos como símbolo para correlacionar ideias diferentes (ARAÚJO, 2006).

Gomes (1996) explica que o método para conceber o esquema de classificação, denominado de Colon Classification, foi o de partir dos assuntos dos livros para identificar as unidades conceituais neles existentes. Estas seriam as peças do quebra-cabeças, que combinadas de diferentes maneiras, poderiam formar qualquer assunto que se desejasse. As unidades foram denominadas de focos e isolados, mas ainda não eram conceitos; ele as chamava de ideias ou perceptos. As unidades poderiam ser conceitos, num momento, mas poderiam ser características destes conceitos, no outro. Segundo Silva (2008, p. 38), “podemos visualizar na Teoria da Classificação Facetada a importância dada aos conceitos, que podem ser considerados como unidades de conhecimento”.

Em 1952, a quarta edição da Colon Classification apresentou pela primeira vez as cinco categorias fundamentais, válidas para qualquer área do conhecimento, conhecidas como PMEST: P (Personalidade), M (Matéria – material ou propriedade), E (Energia- ação), S (Space - local) e T (Time - época) (BARBOSA, 1972).

O Classification Research Group (CRG)², na Inglaterra, desdobrou as categorias de Ranganathan para a elaboração de classificações facetadas, aplicáveis a qualquer temática de tesouros:

- Coisas, substâncias, entidades (produtos, instrumentos, constructos mentais);
- Suas partes constituintes (órgãos);
- Sistemas de coisas;
- Atributos de coisas (qualidades, propriedades, incluindo estrutura, medidas, processo, comportamento);
- Objeto da ação (paciente);
- Relações entre coisas (interações, efeitos, reações);
- Operações sobre coisas (experimentos, ensaios, operações mentais);
- Propriedades de atributos;
- Lugar
- Tempo (CAMPOS e GOMES, 2006).

Tristão (2004) considera que a classificação está presente não apenas nos sistemas que visam à recuperação de informação, mas na base de sistemas e atividades que se ocupam da organização do conhecimento em suas diferentes manifestações, como nos sistemas de inteligência artificial e hipertextos.

4.3 Taxonomia

² Classification Research Group (CRG), criado na Inglaterra em 1952, com o objetivo de desenvolver estudos no âmbito da classificação (CAMPOS e GOMES, 2006).

Terra (2005) define taxonomia como um sistema para classificar e facilitar o acesso à informação, e que tem como objetivos: representar conceitos através de termos; agilizar a comunicação entre especialistas e entre estes 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 processos de conhecimento.

La taxonomía es, en un sentido amplio, la creación de la estructura (orden) y de los rótulos (nombres) que ayudan a localizar la información relevante. En un sentido más específico, es el ordenamiento y rotulación de los metadatos, que permiten gestionar sistemáticamente la información primaria. (MARTINEZ, et al., 2004).

Esse tipo de relação é definida por Dahlberg (1978, p. 104) como sendo aquela que aparece “entre dois conceitos que têm idênticas características, sendo, porém, que uma em relação a outra apresenta uma característica adicional, de modo que surge entre eles uma hierarquia”.

Conforme Campos e Gomes (2008), tradicionalmente, as taxonomias tiveram por função a classificação das espécies em Botânica e Zoologia, adotando uma nomenclatura binária e são, por definição, classificações sistemáticas. Segundo Vital (2011), a palavra 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, de acordo com Vital (2007), apesar das taxonomias serem derivadas da Biologia, nos ambientes digitais, seu aparecimento e uso estão relacionados com as formas automatizadas de organização da informação, tornando-se alvo de estudos da Ciência da Informação e áreas correlatas.

Terra (2005, p. 2) declara que a “importância da taxonomia para informações não estruturadas é apontada, por alguns especialistas, como equivalente à importância que os bancos de dados tiveram para as informações tabulares”.

No âmbito da Ciência da Informação as taxonomias podem ser comparadas a estruturas classificatórias como as Tabelas de Classificação, que têm como objetivo reunir documentos de forma lógica e classificada. Atualmente, as taxonomias reúnem todo tipo de documento digital e permitem, diferentemente das estratégias de busca, um acesso imediato à informação. Ao contrário das tabelas, que oferecem um endereço (notação) que localiza os documentos nas estantes, a taxonomia prescinde de notação (CAMPOS e GOMES, 2008)

Segundo Campos e Gomes (2008), as taxonomias adotam alguns princípios básicos de classificação, como a categorização, os cânones e os princípios para a ordenação das classes e de seus elementos. Para elas, o resgate da taxonomia nos sistemas de informação considera a unidade sistemática (*taxon*) não mais família, gênero, espécie, mas conceitos. Além disso, há uma diversidade nas relações

taxonômicas: “[...] pode-se, atualmente, construir taxonomias policotômicas, ou seja, onde um elemento é associado a tantas classes e subclasses quantas necessárias, dentro de um domínio especializado ou tarefa”.

Vital (2011) afirma que a taxonomia, como qualquer outra forma de representação do conhecimento, precisa acompanhar a natureza dinâmica e as variações linguísticas do ambiente em que foi desenvolvida, e, conseqüentemente, estar em constante avaliação. As taxonomias são eficientes no processo de desenvolvimento das ontologias, instituindo classes e subclasses, com relacionamentos hierárquicos e até mesmo associativos, corroborando no sentido de instituir um mapa semântico de uma determinada área. Porém, não centram seu objetivo na busca de um consenso linguístico em áreas do conhecimento, mas pretendem fazer sentido aos usuários do ambiente no qual foram desenvolvidas, e para isso, utilizam-se de grande flexibilidade em seu desenvolvimento.

4.4 Tesouro

De acordo com Moreira, Alvarenga e Oliveira (2004), o termo "*thesauros*" se origina do grego "*thesaurós*" e significa tesouro ou repositório. Este termo surgiu com a publicação do dicionário analógico de Peter Mark Roget, em Londres, em 1852, intitulado "*Thesaurus of English words and phrases*". O termo "*thesaurus*" também designa vocabulário, dicionário ou léxico, porém o dicionário de Roget se diferenciava dos outros por ser um vocabulário organizado de acordo com o seu significado e não de acordo com a ordem alfabética. A obra teve o mérito de estabelecer a denominação para vocabulários que relacionam seus termos por meio de algum tipo de relação de significado.

Em sua introdução, Roget define seu dicionário como uma “classificação de ideias” e explica que, diferentemente dos outros dicionários, o seu permite que se chegue a uma palavra mais adequada ou que melhor se ajuste às necessidades do escritor, sem que, de início, ele saiba qual é ela (GOMES, 1996).

Nos anos 1960, Vickery apresentou quatro significados para o termo tesouro na literatura da ciência da informação, sendo que o significado mais comum é o de uma lista alfabética de palavras, onde cada palavra é seguida de outras relacionadas a ela (VICKERY, 1980 *apud* MOREIRA, ALVARENGA e OLIVEIRA, 2004).

Currás (1995, p. 88) define tesouro como “uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos linguísticos que a compõem – termos, simples ou compostos – encontram-se relacionados entre si sintática e semanticamente”.

Tristão (2004, p. 167) o define como “vocabulário de termos, que nada mais é do que uma seleção de termos, baseados em análise de conceitos, na qual se define o termo geral, de maior abrangência, e sua relação com termos mais específicos, que representam os conceitos menores”.

Um tesauro é um vocabulário controlado organizado em uma ordem preestabelecida e estruturado de modo que os relacionamentos de equivalência, de homografia, de hierarquia, e de associação entre termos sejam indicados claramente e identificados por indicadores de relacionamento padronizados empregados reciprocamente. As finalidades primordiais de um tesauro são (a) facilitar a recuperação dos documentos e (b) alcançar a consistência na indexação dos documentos escritos ou registrados de outra forma e outros tipos, principalmente para sistemas de armazenamento e de recuperação de informação pós-coordenados (ANSI/NISO Z39.19³, 2003 *apud* SALES e CAFÉ, 2008).

Conforme Campos e Gomes (2006), a evolução na construção de tesauros baseia-se em duas vertentes, a americana e a europeia. Os tesauros elaborados na América do Norte, especificamente nos Estados Unidos, a partir da década de 1950, foram fruto do desenvolvimento que ocorreu a partir do cabeçalho de assuntos para o unitermo e passou de um sistema pré-coordenado para sistemas pós-coordenados.

Silva (2008) afirma que na mesma época, na Inglaterra, o CRG, a partir da Teoria da Classificação Facetada, ampliou as categorias do PMEST e desenvolveu diversas tabelas de classificação, dando origem a uma técnica denominada *Thesaurofacet*. Essa técnica permitiu o melhor posicionamento do conceito dentro do sistema de conceitos em uma dada área de assunto, por meio de uso de suas categorias. A teoria desenvolvida por Ranganathan juntamente com a Teoria do Conceito, de Dahlberg, e princípios terminológicos embasaram também os tesauros terminológicos. Esses instrumentos têm nas características do conceito um elemento essencial para evidenciar as relações entre os conceitos e seu posicionamento no sistema, além de defini-lo.

Dahlberg apresenta, no final da década de 1970, a Teoria Analítica do Conceito voltada para o Referente, mais conhecida como Teoria do Conceito, agregando princípios terminológicos relacionados ao conteúdo conceitual e sua definição. Para Campos e Gomes (2006), estabelece-se assim, uma teoria consolidada para a determinação do que se entenderia por menor unidade em um tesauro, ou seja, o conceito representado por um termo. Além disso, Moreira (2003)

³ AMERICAN NATIONAL STANDARD INSTITUTE. National Information Standards Organization. Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Maryland, 2005.

aponta como inovação, o uso de definições para o posicionamento do conceito no sistema de conceitos.

Bräscher (2010) aponta como função dos tesouros a tradução da linguagem dos documentos, dos indexadores e dos pesquisadores numa linguagem controlada, usada na indexação e na recuperação de informações em sistemas de informação. Conforme Sales e Café (2007), o ANSI/NISO Z39.19 (2003) ressalta que os tesouros não são utilizados somente pelos especialistas da informação no momento da indexação, mas também por usuários da informação no momento da busca de documentos.

4.5 Ontologia

Ontologia como um ramo da Filosofia remete ao “estudo do Ser; a área da metafísica que se relaciona ao Ser ou essência das coisas, ou o Ser no sentido abstrato” (OXFORD English Dictionary *apud* SILVA, 2008).

Na filosofia, uma ontologia é um sistema particular de categorias que versa sobre uma certa visão do mundo (GUARINO, 1998 *apud* MOREIRA e OLIVEIRA, 2005). Sowa (2000) *apud* MOREIRA e OLIVEIRA (2005) define ontologia “como estudo das categorias das coisas que existem ou podem existir em algum domínio”.

Boccatto (2008) aponta que a origem da apropriação de tal termo pela Ciência da Computação remete ao início da década de 1970, quando pesquisas na área da Inteligência Artificial (IA), impulsionaram o desenvolvimento de projetos relacionados à criação e organização de bases de conhecimento computacionais. Na Engenharia de *Software*, as primeiras iniciativas remontam à década de 1980 e na Ciência da Informação (CI) à década de 1990 (ALMEIDA, 2010).

Guarino e Giaretta (1995) *apud* Ramalho (2006) propuseram uma diferenciação entre a ontologia estudada pela área de Ciência da Computação e a Ontologia filosófica de Aristóteles, segundo a qual “uma ontologia” (com o artigo indefinido e inicial minúscula), diz respeito a um determinado objeto em particular, enquanto “Ontologia” (sem o artigo indefinido e com a inicial maiúscula), refere-se à disciplina filosófica que lida com a natureza e a organização da realidade.

Para Schiessl e Bräscher (2012), o termo ontologia, na CC e na CI, tem sido usado de forma ambígua. Conforme a interpretação vai de artefato computacional que armazena e organiza informações até estrutura semântica que representa o conhecimento de um dado domínio.

Em seu uso mais predominante na IA, uma ontologia é referida como um artefato de engenharia, constituído por um vocabulário específico utilizado para

descrever uma certa realidade e um conjunto de pressupostos explícitos relacionados com o significado pretendido para as palavras do vocabulário (GUARINO,1998 *apud* BOCCATO, 2008).

Segundo Moreira e Oliveira (2005), na computação, o termo ontologia denota um sistema de conceitos, da mesma forma que os tesouros, e como tal possui uma natureza epistemológica e não ontológica, ou seja, a preocupação está na estruturação de conceitos e não no entendimento da natureza dos objetos.

Almeida e Bax (2003) afirmam 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”.

É um artefato tecnológico que descreve um modelo conceitual de um determinado domínio em uma linguagem lógica e formal, a partir da descrição dos aspectos semânticos de conteúdos informacionais, possibilitando a realização de inferências automáticas por programas computacionais (RAMALHO, 2006, p. 97).

Dziekaniak (2010) considera que as ontologias são talvez o melhor instrumento para classificação e representação do conhecimento até então pensado e desenvolvido, fornecendo ampla possibilidade de uso inclusive por atores que não dominam a linguagem de máquina.

Souza (2011) acredita que para realizar a tarefa de recuperar informação na *web*, o uso de ontologias é fundamental, pois elas fornecem um vocabulário compartilhado para modelar determinado domínio (área do conhecimento), seus conceitos, propriedades, relações e restrições.

Gruber (1993) *apud* Boccato (20068, p. 203) define ontologia como “uma especificação formal e explícita de uma conceitualização compartilhada”, onde conceitualização representa um modelo abstrato de algum fenômeno que identifica os conceitos relevantes para si mesmo; explícita, significa que os elementos e suas restrições são claramente definidos; formal, na medida em que deve ser passível de processamento automático (por máquinas); e compartilhada, por capturar o conhecimento consensual, aceito por um grupo de pessoas.

Conforme Boccato (2008), apesar de tal definição ser a mais referenciada na literatura relacionada às ontologias, no contexto da representação da informação, Guarino e Giaretta (1995) apontam problemas quanto à noção de conceitualização utilizada por Gruber (1993). Para os autores, a questão está no fato de Gruber considerar uma conceitualização como um conjunto de relações extensíveis descrevendo um estado particular das coisas, enquanto que comumente a noção que se tem de conceitualização é de algo intencional, como uma grade que deve ser preenchida com os vários estados das coisas.

Guarino (1998) *apud* Almeida e Bax (2003, p. 9) define ontologia como:

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. Este conjunto de fatos tem a forma da teoria da lógica de primeira ordem, onde as palavras do vocabulário aparecem como predicados unários ou binários.

Almeida e Bax (2003) explicam que o vocabulário formado por predicados lógicos forma a rede conceitual que confere o caráter intensional às ontologias e que estas definem as regras que regulam a combinação entre os termos e as relações. As relações entre os termos são criadas por especialistas, e os usuários formulam consultas usando os conceitos especificados. Desse modo, uma ontologia define uma “linguagem” usada para formular consultas.

As ontologias são compostas a partir de um conjunto de termos estruturados hierarquicamente, formando, desta maneira, uma estrutura de classes e subclasses. As classes representam as entidades do “mundo real”, que são agrupadas e categorizadas de acordo com suas similaridades, levando-se em consideração um domínio concreto. [...] Uma classe pode possuir diversas subclasses, que podem organizar as entidades a partir de características mais específicas. [...] As entidades descritas em uma ontologia são denominadas como instâncias e consistem de representações de objetos ou indivíduos pertencentes a uma classe ou subclasse. Cada instância possui determinadas propriedades ou atributos (BOCCATO, 2008, p. 203).

Almeida e Bax (2003) citam, ainda, os axiomas, usados para modelar sentenças sempre verdadeiras. São os axiomas que fornecem as informações factuais básicas a partir das quais podem ser derivadas conclusões úteis, por meio de regras de inferências (RUSSEL e NORVING, 2004 *apud* RAMALHO, 2006). Essas regras, de acordo com Campos (2007), fornecem maior formalismo, através da caracterização de aspectos que podem ser utilizados para inferências, como, por exemplo: transitividade, condições necessárias e suficientes, dentre outros.

Sales (2008) destaca que as relações nas ontologias podem se apresentar como relações terminológicas e relações formais. A primeira refere-se à linguagem natural e a segunda, à linguagem compreendida pelas máquinas. No que se refere à parte terminológica, o uso das relações categorias advindas da Ciência da Informação e da Terminologia são bastante úteis. Na parte formal, a Ciência da Computação contribui com as definições formais, inteligíveis para a máquina.

Em relação à finalidade, as ontologias podem:

- compartilhar a compreensão da estrutura da informação entre pessoas ou agentes construtores de softwares;
- permitir a reutilização do conhecimento pertencente a um domínio;
- permitir a representação explícita dos conceitos de um domínio;

- separar o conhecimento de um domínio, do conhecimento que se pode denominar de operacional;
- possibilitar a análise de um domínio (GUZMAN LUNA et al., 2006 *apud* BOCCATO, 2008, p. 204).

Freitas (2004) destaca que as ontologias servem como ferramenta para organização, reuso e disseminação de conhecimento, facilitando a construção de novos agentes e desempenham um papel ainda mais importante, que é o de servir como vocabulário de comunicação entre agentes inteligentes.

As ontologias, quanto aos níveis podem ser:

- a) Ontologia genérica: conceitos muito genéricos, independentes de um problema ou domínio particular;
- b) Ontologia de domínio: descrevem o vocabulário relativo a um domínio específico, através da especialização de conceitos presentes na ontologia de alto nível;
- c) Ontologia de tarefa: descreve o vocabulário pertinente a uma tarefa genérica ou específica através da especialização de conceitos presentes na ontologia de alto nível;
- d) Ontologia de aplicação: descreve conceitos dependentes do domínio e da tarefa particulares (GUARINO, 1994, 1997 *apud* RIOS, 2005, p. 5).

Na visão de Uschold (1996) *apud* Rios (2005), as ontologias podem ser classificadas em três categorias, quanto ao tipo de conhecimento que representam:

- a) Ontologia de tarefas: conceituações sobre a resolução de problemas independentemente do domínio onde ocorram;
- b) Ontologia de representação: conceituações que fundamentam os formalismos de representação do conhecimento;
- c) Ontologia de domínio: conceituações de domínios particulares. Campos (2011) define a ontologia de domínio como a que busca descrever formalmente classes de conceitos e os relacionamentos de determinada área com o objetivo de compartilhar um consenso terminológico.

Quanto ao grau de formalidade, Uschold (1996) as divide em: rigorosamente formal, semiformal e altamente informal. Campos (2007) explica que as primeiras utilizam regras e são projetadas com o intuito de permitirem inferências mais precisas; as ontologias informais são também chamadas de “ontologias leves” e são vocabulários simples, constituídos basicamente por uma hierarquia de termos e relações básicas, do tipo *é-um* e *parte-todo*, e as ontologias semi-formais são um meio termo entre as duas já citadas. Não possuem um formalismo preciso, mas já possuem algumas relações identificadas e um maior rigor na formação das suas hierarquias de conceitos.

Para Almeida e Bax (2003), mesmo sem um consenso, pode-se observar que os tipos apresentados guardam semelhanças entre suas funções e uma vez

conhecidos os principais tipos e características, pode-se buscar ontologias existentes adequadas à utilização desejada.

Almeida (2005) aponta como áreas de aplicação das ontologias, dentre outras: a gestão do conhecimento, o comércio eletrônico, o processamento de linguagens naturais, a recuperação da informação na *Web*. Atualmente, as ontologias são uma recomendação do W3C⁴ para a Web Semântica⁵ como padrão de vocabulário comum para a troca de dados, para tornar o conhecimento reutilizável e para facilitar a comunicação de sistemas heterogêneos (CAMILO e SILVA, 2009).

Tratando-se especificamente de ontologia de domínio, considera-se que um dos seus propósitos é minimizar a heterogeneidade conceitual através da representação formal de conceitos e relacionamentos de um domínio de discurso (PADILHA, BAIÃO e REVOREDO, 2012).

De acordo com Guizzardi et al. (2011), uma ontologia de domínio é um tipo particular de modelo conceitual que deve satisfazer o requisito de servir como uma representação de consenso de uma conceituação compartilhada por uma determinada comunidade. Modelos conceituais, por sua vez, são definidos por Padilha, Baião e Revoredo (2012, p. 8) como “artefatos produzidos com o objetivo de representar uma porção da realidade segundo uma determinada conceituação”.

As ontologias de domínio são o tipo de ontologia mais comumente desenvolvido e para Silva (2008) elas necessitam de ontologias fundamentação para a concepção da sua estrutura inicial, no que concorda Santos (2010), quando afirma que os termos e as propriedades de uma ontologia de domínio são obtidos através da especialização de conceitos de uma ontologia de topo⁶.

Guizzardi et al.(2011) defendem que se uma ontologia de domínio é um modelo conceitual, uma linguagem adequada para representação de ontologias de domínio deve satisfazer os requisitos gerais de uma linguagem adequada para modelagem conceitual, ou seja, deve ter como teoria subjacente uma ontologia de fundamentação.

Uma ontologia de fundamentação apresenta princípios que estão concernentes com uma ontologia formal. Esses princípios, independentes de um domínio, permitem a elaboração de modelos para a representação de diversos contextos de representação, sendo altamente reutilizáveis. É também caracterizada por ser filosoficamente bem fundamentada, permitindo a explicitação de uma

⁴ O Consórcio da Web ou W3C (*World-Wide Web Consortium*)

⁵ Também conhecida como Web 3.0, é uma extensão da Web atual, na qual é dada à informação um significado semanticamente definido, permitindo que computadores e pessoas trabalhem em cooperação (BERNERS-LEE et al., 2001 *apud* BRAVO, 2010).

⁶ O mesmo que Ontologia de Fundamentação

visão da realidade, ou seja, do acordo ontológico estabelecido, com determinação de regras de restrição, bem como conceitos, categorias e metapropriedades (MEDEIROS e CAMPOS, 2005, p. 214).

4.6 Modelagem conceitual

Cougo (1997) *apud* Campos (2003, p. 9) define modelo como “a representação abstrata e simplificada de um sistema real, com a qual se pode explicar ou testar seu comportamento, em seu todo ou em partes”.

Le Moigne afirma que conhecer é modelizar, ou seja, o processo de conhecer equivale à construção de modelos do mundo/domínio a ser construído que permitem descrever e fornecer explicações sobre os fenômenos que observamos (CAMPOS, 2004).

Modelos são representações simplificadas da realidade que se quer compreender. [...] O mundo é complexo e modelos são produzidos para que se possa compreendê-lo em parte, visto que não se consegue abrangê-lo em sua totalidade (TEIXEIRA, 2009).

Segundo Medeiros e Campos (2005), é importante que a construção de sistemas de recuperação da informação sejam baseados em modelos construídos sobre uma realidade. Esses modelos são representações parciais de determinado mundo, onde é possível representar a existência de objetos e as relações entre eles.

Campos (2004) evidencia que, antes da utilização do modelo, é necessário ter bastante claro o método utilizado para a sua construção e acredita que tais métodos não são possíveis de serem pensados de forma separada, mas sim de forma sistêmica, segundo a qual o indutivo e o dedutivo, a análise e a síntese, são vistos como processos não disjuntos, que se complementam em um todo, ou seja, “análisesínteseanálisesíntese...”.

Conceituar é “criar, desenvolver e/ou enunciar conceito acerca de; definir, conceitualizar, conceptualizar” (DICIONÁRIO Eletrônico Houaiss, 2004).

Conceito pode ser definido como:

processo que torne possível a descrição, a classificação e a previsão dos objetos cognoscíveis. Assim entendido, esse termo tem significado generalíssimo e pode incluir qualquer espécie de sinal ou procedimento semântico, seja qual for o objeto a que se refere, abstrato ou concreto, próximo ou distante, universal ou individual (ABBAGNANO, 1998, p. 167).

Na Ciência da Informação, Dahlberg (1978a, p. 12) define conceito como a “unidade do conhecimento”.

Segundo Gonçalves e Souza (2009), “conceito” em ontologia compõe desde as classes mais gerais até as instâncias. Mesmo as relações (propriedades e

atributos das classes) são basicamente conceitos que caracterizam e completam o sistema terminológico da ontologia.

A norma ISO 704 (1987) dispõe que conceitos são “construções mentais que servem para classificar os objetos individuais do mundo exterior ou interior através de um processo de abstração mais ou menos arbitrário”.

Um conceito é uma rede de padrões de inferências, associações e relacionamentos que são predicados ou ditos de outra forma trazidos em cena através do ato da categorização [...] a cristalização ou formalização do pensamento inferencial, nascida da percepção sensorial, condicionada pela operação do cérebro humano e delineada pela experiência humana. Ela repousa na fundamentação de todo pensamento, mas ela é pragmática e instrumental. É permanente e efêmera. Permanente, porque sem ela, a cognição é impossível; efêmera porque ela pode ser rejeitada quando sua utilidade é esgotada (SHERA, 1957 *apud* ALVARENGA, 2001, p. 33).

Retomando a definição de Gruber (1993) de que uma ontologia é uma especificação explícita de uma conceitualização⁷, tem-se que o termo conceitualização corresponde a uma coleção de objetos, conceitos e outras entidades que se assume existir em um domínio e os relacionamentos entre eles. Uma conceitualização é uma visão abstrata e simplificada do mundo que se deseja representar (ALMEIDA, 2005).

Guarino e Welty (2009) alertam para o fato de que uma conceitualização trata de conceitos. Ontologia, na sua essência, deve tratar de um conceito independente do seu estado no mundo, isto é, da intensão de um conceito que Dahlberg (1978) sugere como a soma total das suas características (SCHIESSL e BRÄSHER, 2012).

Diversos autores entendem que a atividade de conceitualização corresponde à criação de um modelo conceitual, em que se procura descrever a realidade, com fins diversos (WAND e WEBER, 1990; GUARINO, 1998 *apud* TEIXEIRA, 2009).

Teixeira (2009) afirma que o processo de conceitualização diz respeito à criação de um modelo conceitual, quando o modelador faz abstrações ao tentar representar a parte da realidade que lhe interessa. Ao criar modelos é preciso privilegiar características mais importantes, que podem variar de um grupo para outro. Assim, podem existir diferentes modelos para representar a mesma realidade. O

⁷ Segundo Almeida (2006), o termo “conceitualização” tem sido adotado, apesar de não existir no português, por sua especificidade na área de Inteligência Artificial. O termo mais próximo no português, “conceituação”, definido no Dicionário Aurélio como “ato ou efeito de conceituar”, não corresponde à definição do termo conceitualização introduzido por GENESERETH e NILSSON (1997): [...] a coleção de entidades que se assume existir em alguma área de interesse e os relacionamentos entre elas (MARANGON, 2009).

processo de conceitualização, aqui entendido como o grupo de processos de abstração que podem resultar em modelos, é abordado por diversos campos de pesquisa: a Filosofia, a CI, a Engenharia de *Software*, dentre outros.

A construção de modelos conceituais está diretamente ligada à representação do conhecimento e estes devem ser capazes de representar um contexto, sendo construídos a partir de processos que evitem qualquer tipo de ambiguidade, ressaltando objetos relevantes ao domínio, bem como seus relacionamentos e atributos (MEDEIROS e CAMPOS, 2005).

No desenvolvimento de sistemas computacionais, segundo Moreira e Oliveira (2005), a modelagem conceitual é a etapa do processo que tem por objetivo “criar um modelo do sistema a desenvolver. Este modelo deve focalizar o que o sistema deverá fazer sem se preocupar com a forma de sua implementação”.

Nos anos 1970 surgiram os primeiros modelos semânticos utilizados em modelagem conceitual em sistemas de gerenciamento de bancos de dados. Almeida, Souza e Coelho (2009) destacam o modelo de dados semânticos, a arquitetura de três esquemas, o modelo Entidade Relacionamento (ER), dentre outros. A principal característica dos modelos semânticos, em relação aos modelos lógicos, é a facilidade de entendimento.

Na área de Inteligência Artificial não existe um diagrama padrão para a modelagem. Vários tipos de representações podem ser usados, tais como diagramas hierárquicos, grafos conceituais, quadros, sendo que em muitos casos, a representação conceitual é também a linguagem de implementação (MOREIRA e OLIVEIRA, 2005).

A Orientação a Objeto é uma abordagem utilizada para a concepção de modelos objetivando a construção de sistemas de informação. A principal técnica para construir modelos orientados a objetos é a *Unified Modeling Language* (UML), que surgiu a partir da compilação de práticas de Engenharia de *Software* na modelagem de sistemas complexos (TEIXEIRA, 2009).

Os modelos semânticos dispõem de um conjunto limitado de constructos para criar um modelo conceitual. O modelo ER, por exemplo, pressupõe que a realidade de interesse para o sistema pode ser representadas por dois constructos apenas (entidade e relacionamento). Segundo Smith e Welty (2001) *apud* Almeida, Souza e Coelho (2009), a inconsistência nas atividades de modelagem das primeiras iniciativas de modelagem conceitual é responsável pelas dificuldades atuais de

interoperabilidade em Sistemas de Informação (SI)⁸. Uma alternativa para solucionar esse tipo de problema são os modelos baseados em ontologias.

4.6.1 Modelagem conceitual em ontologias

As ontologias são planejadas e construídas pelo homem para atender suas necessidades de informação em relação à organização, à estruturação e à recuperação. Dessa forma, elas são parte integrante de um sistema formal que de alguma maneira necessita de uma forma de representação (SCHIESSL e BRÄSHER, 2012).

O estudo do ciclo de vida da ontologia identifica um conjunto de estágios através dos quais o processo de desenvolvimento é conduzido. Fernandez-López et al (1997) *apud* Teixeira (2009), indica que esses estágios são: especificação, conceitualização, formalização, integração, implementação e manutenção.

No desenvolvimento de uma ontologia, a fase de conceitualização é a mais crítica do desenvolvimento, pois, sendo a mais ligada à definição do conhecimento, nela ocorrem a maior parte das atividades de suporte de aquisição e avaliação (FREITAS, 2004).

Segundo Almeida (2006) e Teixeira (2009), a conceitualização é uma atividade intelectual e ferramentas automatizadas não podem prover auxílio significativo.

Nessa fase, a estrutura do domínio de conhecimento é traduzida em um modelo conceitual. Conforme Teixeira (2009), tal modelo descreve soluções relativas ao vocabulário do domínio identificado na atividade de especificação. Constrói-se um glossário de termos, incluindo conceitos, verbos, instâncias e propriedades, o qual identifica e reúne o conhecimento do domínio. Na sequência agrupam-se termos, como substantivos e verbos, que representam conceitos e relações, respectivamente, na ontologia.

No processo de modelagem, é preciso compreender a linguagem utilizada no universo do discurso e depois representar esses conceitos em uma linguagem de modelagem conceitual; para tal, faz-se uso de fontes de informações em uma linguagem natural, que podem incluir documentos oficiais, descrições de outros sistemas de informação, entrevistas, ou até mesmo simples observação (BAIÃO, 2009).

⁸ SI – sistema no qual a entrada principal são os dados, que através da seleção, tratamento e organização são transformados em informação para algum usuário (SILVA, 2008).

Kabilan (2007) *apud* Teixeira (2009) sugere a análise do vocabulário tipicamente obtido em um processo de conceitualização, através da pluralidade de perspectivas e de visões como: o domínio do conhecimento através da extração, análise e representação dos conceitos e de seus significados (semântica), suas intenções (prática) e seu plano de ação (conhecimento dos processos).

Dentre as metodologias de modelagem conceitual em ontologias, Almeida (2006) considera a *Methontology* a metodologia mais madura para essa finalidade, em relação ao padrão de desenvolvimento de *software* (IEEE-1074 67)⁹.

Para Freitas (2004), a comunidade de desenvolvimento de sistemas orientados a objetos têm despertado para o uso de ontologias através da linguagem diagramática de modelagem unificada UML. Essa linguagem não pode ser considerada um formalismo de representação, entretanto, possui construtos abstratos o suficiente para permitirem a representação de ontologias, como classes e atributos.

Guizzardi (2005) defende que a *Unified Foundational Ontology* (UFO)¹⁰ é relevante no desenvolvimento de ontologias na medida em que a associação das entidades do modelo às classes de fundamentação explicita meta-propriedades apoiadas em bases teóricas sólidas (PADILHA, 2012).

Desde o fim da década de 1980, observa-se um crescente interesse no uso das ontologias de fundamentação no processo de avaliação e (re)engenharia de linguagens de modelagem conceitual. Considerando que modelos conceituais são artefatos produzidos com o objetivo de representar uma determinada porção da realidade segundo uma determinada conceituação e que Ontologias de Fundamentação descrevem as categorias que são usadas para a construção dessas conceituações, Guizzardi et al. (2011) concluem que uma linguagem adequada de modelagem conceitual deveria possuir primitivas de modelagem que refletissem as categorias conceituais definidas em uma Ontologia de Fundamentação.

⁹Software Engineering Standards Committee of the IEEE Computer Society. IEEE Standard for Developing Software Life Cycle Processes. December 1997.

¹⁰ Guizzardi (2005) apresenta a construção de uma ontologia de fundamentação, denominada *Unified Foundational Ontology* (UFO), para o apoio a modelagem conceitual. Essa ontologia é dividida em três fragmentos: 1) UFO-A, é a ontologia de *endurants*² (*endurantes*), que define o cerne da UFO (com termos como, por exemplo, tipos, relações todo-parte, espaço de valores de atributos, entre outros); 2) UFO-B define, adicionalmente à UFO-A, uma ontologia com termos relativos aos *perdurants* (ou *perdurantes*, como, por exemplo, estados, processos, eventos, relações temporais, entre outros.); e 3) UFO-C define, adicionalmente à UFO-B, termos relacionados a conceitos sociais (tais como plano, ação, objetivo, agente, intencionalidade, comprometimento e compromisso).

4.7 Fundamentos da Ciência da Informação aplicados à construção de ontologias

No início do século XX, o termo “Documentação” foi cunhado por Paul Otlet, responsável também por sua sistematização. Em sua obra *Traité de Documentation* (1934), mostra-se interessado em toda novidade tecnológica capaz de condensar e organizar a informação de acordo com suas necessidades e objetivos. Segundo Ferneda (2003), Otlet juntamente com Henri La Fontaine funda o *International Institute of Bibliography*, marco do que veio a se chamar Documentação e posteriormente Ciência da Informação. Shera e Cleveland (1977) *apud* Ferneda (2003) afirmam que a década de 1960 forneceu um clima favorável ao desenvolvimento da CI, pois problemas relacionados ao tratamento da informação começavam a ser abordados pela comunidade científica mundial, ao mesmo tempo em que se vivia um período de acelerado desenvolvimento tecnológico.

A natureza interdisciplinar da CI propicia o surgimento de diferentes correntes e estimula discussões sobre o seu objeto de estudo, a informação (FERNEDA, 2003).

Ferneda (2003) afirma que na Ciência da Computação, a definição de informação se refere mais à construção de sistemas informáticos nos quais dados podem ser totalmente descritos através de representações formais, podendo ser quantificados, armazenados e processados pelo computador. Já na CI, a informação está diretamente relacionada ao seu significado o que implica procedimentos menos formais ou operacionais baseados na capacidade e na habilidade de abstração, apreensão e representação da significação, contextualizando-a. Para o autor esta operação intelectual não pode ser realizada de forma absoluta por modelos computacionais, pois estes trabalham apenas com formas significantes.

A Ciência da Informação pode ser entendida, na visão de Saracevic (1996, p. 47):

como um campo dedicado às pesquisas científicas e à prática profissional voltadas para os problemas da efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação.

Enquanto a Ciência da Informação é um campo dedicado a abordar os problemas de comunicação de conhecimento e de seus registros entre seres humanos (Saracevic, 1996), a Ciência da Computação enfatiza a busca por melhorias nos processos algorítmicos (ALMEIDA, 2010).

Na CI, a Recuperação da Informação se firmou como área de pesquisa e para Saracevic (1996) é a vertente tecnológica da CI e resultado da integração desta com a Ciência da Computação.

O objetivo distinto dessas áreas de pesquisa reflete em seus profissionais, em suas habilidades e em suas possibilidades de atuação, no âmbito de projetos de Sistemas de Informação (SI). A formação em computação enfatiza a codificação e os procedimentos técnicos da atividade de desenvolvimento. Os profissionais da CI, em geral recebem fundamentação sobre a natureza da informação e do conhecimento, estudam teorias sobre organização da informação, princípios de classificação, fundamentos dos instrumentos de organização (como por exemplo, vocabulários controlados), estudam necessidades de usuários, dentre outros assuntos (ALMEIDA, 2010, p. 49).

A CI apresenta similaridades entre instrumentos utilizados na Biblioteconomia e na Inteligência Artificial. Elas estão na forma de elaboração da estrutura de instrumentos que demanda a organização de conceitos em processos que incluem categorização e classificação de conceitos, definição das relações entre eles e tratamento da terminologia empregada nos conceitos e relações (SILVA, 2008).

Ferneda (2003) considera que a web semântica utiliza procedimentos da CI como: definição de formatos de intercâmbio de registros bibliográficos e desenvolvimento de tesouros. E complementa afirmando que no projeto da web semântica estão inseridos conceitos e ideias que há muito tempo são utilizados na CI no tratamento documental. O autor alerta que muitos trabalhos referentes ao tratamento de informação textual no âmbito da Ciência da Computação ignoram a existência da Ciência da Informação, e desse modo propõem métodos e técnicas há muito utilizados pela CI.

Na atividade de desenvolvimento de sistemas de informação são criados modelos conceituais que, segundo Almeida e Barbosa (2009), são obtidos a partir de abstrações de aspectos da realidade, na perspectiva de uma pessoa ou grupo de pessoas. Para Almeida (2010), a modelagem inconsistente resulta em SI com várias deficiências, as quais muitas vezes se tornam aparentes apenas em longo prazo. Tal prática se coloca no sentido oposto a todo o esforço despendido para melhor organizar e manipular a informação e o conhecimento. Problemas verificados nos SI das instituições modernas, como a dificuldade de integração, têm origem em atividades de modelagem deficientes, conduzidas de forma *ad-hoc* e sem correspondência com o mundo real (SMITH e WELTY, 2001; GUARINO, 1998 *apud* ALMEIDA, 2010).

Almeida (2010) aponta as habilidades do profissional de CI como essenciais na criação de modelos que representem a realidade, de forma que se possa obter SI eficientes quanto aos seus objetivos no contexto social em que atuam.

Teixeira (2009) ressalta que o esforço para caracterizar ontologias como parte integrante da Ciência da Informação é relevante e necessário uma vez que seus estudos originaram-se no âmbito da Ciência da Computação, mas contribuições significativas têm sido originadas na CI e enriquecem e caracterizam de forma interdisciplinar o assunto.

Para Barquín, González e Pinto (2006), as taxonomias e os tesouros podem atuar como núcleo sobre o qual se construirá a ontologia. Em consequência haverá um vocabulário controlado com relações típicas das taxonomias e tesouros, além de propriedades, relações e inferências. A parte da conceitualização das ontologias, isto é, o vocabulário de termos controlados é equivalente ao que já se conhece como tesouro (MOREIRO GONZÁLEZ, 2004 *apud* BARQUÍN, GONZÁLEZ e PINTO, 2006).

A necessidade de explicação e relacionamento entre os termos aproximam as ontologias das estruturas de construção de linguagens documentárias, pois:

uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõe uma estrutura no domínio e restringe as possíveis interpretações dos termos (JASPER e USCHOLD, 1999 *apud* BOCCATO, 2008, p. 203)

Nesta perspectiva, de acordo com Ramalho (2006), ao longo dos últimos anos inúmeros estudos comparativos entre tesouros e ontologias têm sido apresentados, devido ao fato de ambos possuírem características semelhantes, como serem constituídas por um conjunto de conceitos e seus respectivos relacionamentos semânticos, representarem recursos informacionais e favorecerem sua posterior recuperação. Contudo, são considerados diferentes modelos de representação do conhecimento. Um fator determinante para a distinção de tais modelos de representação é a necessidade de, nas ontologias, descrever-se formalmente todas as relações e para se rotular tais relacionamentos normalmente são utilizados verbos ou verbos preposicionados. Além disso, o autor ainda aponta como outra característica das ontologias, a possibilidade de representarem tanto informações genéricas como informações concretas, denominadas como instâncias.

Moreira (2003) ressalta que os tesouros são instrumentos terminológicos e por isso registram parcialmente os comprometimentos ontológicos, no entanto, enquadrá-los como um tipo de ontologia pode ser precipitado, uma vez que existem diferentes princípios envolvendo a construção de cada instrumento. Keizer et al. (2000) *apud* Moreira (2003) também não enquadram tesouros como um tipo de ontologia. Eles consideram tesouros, classificações, vocabulários, nomenclaturas e sistemas de codificação sob a denominação de sistemas terminológicos, uma vez que o foco das

ontologias é mais direcionado para os conceitos, enquanto que, nos sistemas terminológicos, o foco é mais direcionado para os termos.

Gonçalves e Souza (2009) consideram como uma das mais importantes contribuições da CI para a construção de ontologias as teorias para a análise conceitual. A partir dos conceitos definem-se as relações envolvidas e, ao final, uma taxonomia subsidia a formalização dessa estrutura, ao ser expressa em linguagem passível de “entendimento” pelo computador.

Na perspectiva da Ciência da Informação, Campos (2001) apresenta cinco diferentes possibilidades de modelização: a) a Teoria da Classificação Facetada; b) a Teoria da Terminologia; c) Teoria do Conceito; d) Orientação a Objetos e e) Ontologias.

As três primeiras possibilidades estão no âmbito da Ciência da Informação, em que são analisadas as teorias ligadas à representação de sistemas de conceitos (CAMPOS, 2004).

A **Teoria da Classificação Facetada de Ranganathan** propõe a categorização do conhecimento em cinco categorias fundamentais, denominadas PMEST, sigla que representa os elementos: personalidade, matéria, energia, espaço e tempo. Faceta é o termo genérico usado para representar algum componente, o qual pode ser um assunto básico ou um assunto isolado. A análise facetada é usada para a construção de tesouros e tabelas de classificação, em que o assunto básico é aquele mais abrangente de um domínio do conhecimento; e a ideia isolada é um conceito dentro deste domínio. A teoria de Ranganathan não lida com relações entre categorias, mas evidencia estruturas hierárquicas.

A **Teoria da Terminologia** proposta por Wüster objetiva determinar o termo que representa um conceito específico (MOREIRA, 2005). Nasceu da necessidade de padronizar termos para a comunicação entre especialistas. A principal preocupação é “discutir a natureza dos conceitos, suas características e relações mútuas, sua descrição (definição), a formação de termos, normalização e internacionalização” (LARA, 2005). A Teoria da Terminologia faz distinção entre objetivos teóricos, os quais oferecem princípios para delimitação de termos e conceitos; e objetivos práticos, relacionado à criação de vocabulários, glossários e dicionários técnicos.

A **Teoria do Conceito** de Dahlberg estende o número de categorias de Ranganathan e também é hierárquica. Trata de categorias para entendimento de um conceito e não da classificação de entidades em um domínio de conhecimento. Os princípios da Teoria do Conceito auxiliam na determinação de um conceito e de suas relações, seja para elaborar tabelas de classificação, seja para construir tesouros. Dahlberg propõe que o conceito seja uma unidade de conhecimento, ao invés de unidade de pensamento, pressupondo assim maior entendimento sobre a área de interesse (TEIXEIRA, 2009, p. 51).

Ainda dentro das teorias da classificação, pode-se posicionar aqui, de acordo com Rivier (1992), a Teoria dos Níveis Integrativos, do CRG, que propôs uma classificação com base naturalista, que refletia a ordem natural do mundo real. Para o grupo, todas as áreas do conhecimento poderiam ser ordenadas pelo princípio de

níveis sucessíveis de complexidade crescente, em que cada nível integra o seu precedente. Essa teoria inspirou a noção de contextualização e para Spiteri (1995) validou o método para abordagem indutiva na classificação. Contudo, a autora afirma que o CRG nunca foi capaz de resolver especificamente como tal teoria poderia ser usada em um sistema de classificação, uma vez que a sua aplicação não responde a uma série de questões e apresenta falhas.

No que diz respeito ao método de raciocínio, de acordo com Moreira e Campos (2005), a Teoria da Classificação Facetada visa o estabelecimento de categorias gerais a partir do olhar sobre um domínio, deixando a compreensão dos objetos que as constituem para um momento posterior, utiliza, portanto o método dedutivo. A Teoria do Conceito, conforme Campos (2004, p. 26), compreende um modo analítico-sintético de conhecer o domínio, sendo uma metodologia “híbrida – não só o método dedutivo e não só o método indutivo –, mas agregando os dois em um exercício de pensar o particular como um todo e o todo possuindo particulares”. Deste modo, entende-se que o domínio apresentará categorias ao final da análise dos conceitos, embora não se conheça inicialmente quais são, chegando-se a elas a partir da análise dos conceitos.

A Ciência da Informação, segundo Medeiros e Campos (2005), dispõe de bases teóricas e metodológicas próprias para a construção de instrumentos terminológicos que constitui um arcabouço sólido de conhecimentos, capaz de permitir que seja criada uma teoria independente sobre um domínio. Este arcabouço está posto na Teoria do Conceito, a qual permite perceber o domínio a partir de uma análise analítico-sintética, e na Teoria da Classificação Facetada, que estabelece categorias gerais de domínio, bem como regras para que isso seja feito de forma válida.

4.8 Teoria do Conceito

Na Teoria do Conceito, Ingetraut Dahlberg (1978a, p. 12) define conceito como “unidade de conhecimento que surge pela síntese dos predicados necessários relacionados com determinado objeto e que, por meio de sinais linguísticos, pode ser comunicado”. A formação dos conceitos é a síntese dos predicados¹¹ necessários verdadeiros a respeito de determinado objeto e para que tal síntese possa ser utilizada no processo da comunicação é necessário acrescentar a ele um nome, uma etiqueta

¹¹ Predicados são as características de determinado objeto. Estas características não devem ser confundidas com as propriedades do conceito, como, por exemplo, a propriedade de pertencer a determinada categoria, de ter relações, de possuir intensão e extensão, etc. (DALHBERG, 1978a).

ou um sinal (DAHLBERG, 1978b). Sem que se tenha efetuado alguma fixação por meio dos sinais linguísticos, nada se poderá relacionar com a mencionada síntese, não podendo o conceito ser utilizado na linguagem (DALHBERG, 1978a).

Segundo Gomes e Campos (2004), Dahlberg considera o conceito como constituído de três elementos: o objeto/o referente, as características desse objeto/referente e a forma verbal que designa esse objeto/referente. O conceito não prescinde destes três elementos: faltando um deles não se tem o conceito, tem-se uma 'ideia' mais ou menos vaga.

O Triângulo do Conceito apresenta:

uma tríade formada por (A) um referente (qualquer objeto material ou imaterial, atividade, propriedade, dimensão, tópico, fato, etc.), (B) os enunciados (predicações) verdadeiros e essenciais sobre um referente que estabelecem as características sobre o referente e (C) o termo, que é forma externa e comunicável do referente e suas características (DAHLBERG, 1981 *apud* MOREIRA, 2003, p. 17).

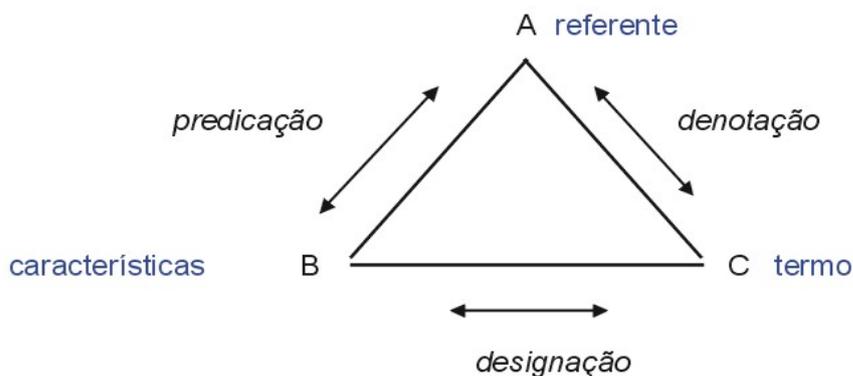


Figura 2: Triângulo do conceito, a representação formal da unidade de conhecimento. Fonte: MARRONI, Gilza Núria Brandão. Identificação e delimitação de relações associativas em tesauros: um estudo de caso na área de Direito do Trabalho. Brasília: UnB, 2006.

Motta (1987) explica que o conceito está contido a um só tempo, no referente, nas predicações (características) e na forma verbal que o designa, ou seja, o conceito é uno. O conceito está para o plano das ideias, assim como o termo está para o plano verbal.

No triângulo proposto por Dahlberg, o conceito é o próprio triângulo. Para Moreira (2003), Dahlberg adota uma visão conceitualista, que reúne as noções de conceito e termo como aspectos de uma mesma entidade. Esta é uma visão adequada à Ciência da Informação, uma vez que lida com linguagens especializadas, onde é preciso exigir uma correspondência unívoca entre termo e conceito. Dahlberg (1978b) afirma que se tratando da comunicação do dia a dia, a imprecisão pode não acarretar grandes consequências, porém, quando se trata de linguagens especializadas as

consequências podem ser desagradáveis. Neste caso, deve-se fazer todo esforço para que os conceitos sejam definidos com toda precisão.

Segundo Moreira (2003), Dahlberg distingue três tipos de conceitos: os conceitos gerais, os conceitos especiais e os conceitos individuais. Os conceitos gerais são aqueles que apontam para todos os itens de um dado tipo (gêneros); os conceitos especiais são aqueles que apontam para alguns itens de um dado tipo (espécie) e, finalmente, os conceitos individuais são aqueles que apontam para um único item de um dado tipo (indivíduo).

Dahlberg também distingue, de acordo com Medeiros (2009), as relações entre conceitos em quantitativas e qualitativas. A primeira (quantitativa) busca comparar dois conceitos distintos de acordo com uma visão formal. Assim, as relações podem ser identificadas da seguinte maneira:

identidade	$A(x, x, x) B(x, x, x)$	As características são as mesmas;
implicação	$A(x, x) B(x, x, x)$	O conceito A está contido no conceito B;
intersecção	$A(x, x, o) B(x, o, o)$	Os dois conceitos coincidem algum elemento;
disjunção	$A(x, x, x) B(o, o, o)$	Os conceitos se excluem mutuamente. Nenhuma característica em comum;
negação	$A(x, x, o) B(o, x, o)$	O conceito A inclui uma característica cuja negação se encontra em B.

Figura 3: Comparação de conceitos do ponto de vista da lógica formal. Fonte: DAHLBERG, Ingetraut, Teoria do conceito. Ciência da Informação, v.7, n. 2, p.101-107, 1978b.

A segunda (qualitativa) pode ser dividida a partir de seus relacionamentos em: formal/categorial, material-paradigmática e funcional-sintagmática, e podem ser assim descritas:

a) Relação formal/categorial: depende do referente conceitual escolhido e é baseada no processo de categorização do mesmo. De acordo com Silva (2008) as categorias e subcategorias formais de Dahlberg são:

Categorias	Subcategorias
Entidades	{ Fenômenos Objetos imateriais Objetos materiais
Propriedades	{ Quantidades Qualidades Relações
Atividades	{ Operações Estados Processos
Dimensões	{ Tempo Posições Espaço

Figura 4: Categorias e subcategorias de Dahlberg. Fonte: MOREIRA, Alexandra. Tesouros e ontologias: estudo de definições presentes na literatura das áreas das ciências da computação e da informação, utilizando-se o método analítico-sintético. Belo horizonte, 2003.

b) Relação material-paradigmática: depende da categoria fundamental do objeto do conceito. Pode ser subdivida em:

I. Relação hierárquica:

- quando dois conceitos apresentam características iguais, sendo que um dos conceitos possui uma ou mais características a mais ou a menos.

Ex.: árvore

 árvore frutífera
 macieira

- relação hierárquica que existe entre os conceitos específicos do mesmo gênero. (tais relações merecem mais o nome de relações coordenadas ou de coordenação — relations in array).

Ex.: árvore

 árvore frutífera
 macieira
 pereira
 pessegueiro
 árvore de nozes
 amendoeira
 aveleira
 noqueira

II. Relação partitiva: relação entre o inteiro e suas partes.

Ex.: árvore

raízes, tronco, galhos, folhas, flores, frutos

III. Relação de oposição: relação de contrariedade entre os conceitos: a oposição entre os conceitos pode ser: contraditoriedade e contrariedade. No primeiro caso, só existem dois momentos, ex.: branco/não branco. No segundo, apresenta-se uma terceira possibilidade: branco/verde (DAHLBERG, 1978b).

c) Relação funcional-sintagmática: aqui estão inseridas as relações associativas. Estas relações aparecem quase exclusivamente na dependência do conceito de processo, ou seja, quando do conceito de processo deriva uma função a ele inerente.

Por exemplo: Pintura (tem como consequência a existência de) quadros (que, por sua vez, supõe um) pintor (assim como de) críticos de arte (ou mesmo de) compradores de quadros, etc. Pode-se dizer que os conceitos de processos possuem determinadas valências que necessariamente deverão ser mencionadas e que as relações emergentes entre tais cadeias de conceitos podem ser reduzidas ao número e espécie dos degraus verificados no processo. Este tipo de relações tem valor não só para a estrutura dos sistemas de conceitos, mas também, já que podem aparecer em qualquer estrutura complexa de predicados, para o processo de definição dos conceitos e principalmente para o processo de análise/síntese dos conteúdos temáticos, ou, mais simplesmente, dos temas a serem tratados em publicações científicas (DAHLBERG, 1978^a, p. 15).

Dahlberg (1978b) ressalta que é fácil verificar que as relações abstrativas ou hierárquicas e as relações partitivas aplicam-se principalmente a conceitos que expressam objetos e que as relações de oposição se aplicam principalmente a conceitos que expressam propriedades, e as relações funcionais, por sua vez, aplicam-se, sobretudo, aos conceitos que expressam processos.

Dahlberg (1978b) trata ainda da intensão e extensão dos conceitos. A intensão do conceito é a soma total das suas características. É também a soma total dos respectivos conceitos genéricos e das diferenças específicas ou características especificadoras. A intensão no conceito de “casa” pode ser representada por:

- edifício;
- habitualmente feito de pedra ou madeira;
- contendo quartos e salas;
- contendo portas e janelas;
- contendo teto, etc.

A extensão do conceito, por sua vez, pode ser entendida como a soma total dos conceitos mais específicos que possui. Pode também ser entendida como a soma dos conceitos para os quais a intensão é verdadeira, ou seja, a classe dos conceitos de tais objetos dos quais se pode afirmar que possuem aquelas

características em comum que se encontram na intensão do mesmo conceito. Distinguem-se duas espécies de extensão:

- a) extensão de um conceito genérico em relação com os conceitos específicos. Ex.: casa - casa de pedra; casa de madeira;
- b) extensão dos possíveis conceitos individuais. Compreende os indivíduos para os quais é válida a predicação genérica do conceito. Ex.: casa - casa do Presidente da República; casa do vizinho, etc.

Dahlberg (1978b) explica que o processo de decomposição do conceito, coletando-se os enunciados verdadeiros sobre determinado objeto pode ser considerado como análise do conceito. Pode-se então dizer que os elementos do conceito são obtidos pelo método analítico-sintético. Cada enunciado apresenta um atributo predicável do objeto que, no nível de conceito, se chama característica. Muitas vezes não se trata de um atributo a que corresponde uma característica, mas de uma hierarquia de características, já que o predicado de um enunciado pode tornar-se sujeito de novo enunciado e assim sucessivamente até atingirmos uma característica tão geral que possa ser considerada uma categoria (entende-se aqui por categoria o conceito na sua mais ampla extensão).

Dahlberg utiliza a noção de categoria sob dois aspectos: como um recurso para o entendimento da natureza do conceito e para a formação de estruturas conceituais. A importância fundamental da categoria na estruturação do conceito e do sistema de conceitos é enfatizada pela mesma autora, quando afirma que se pode ver que as categorias têm uma capacidade de estrutura: não apenas estruturam todos os elementos de conhecimento e unidades de conhecimento; elas fornecem, ao mesmo tempo, por esse meio o esqueleto, os ossos e os tendões para estruturar todo o conhecimento. Com o uso consciencioso, então, o corpo do conhecimento pode se manter unido, pode se mover, pode se manter flexível e pode crescer organicamente (TRISTÃO, 2004, p. 168).

Dahlberg, conforme Campos e Gomes (2006), desenvolve sua teoria que fornece bases seguras para estruturar as hierarquias (relações lógicas). Posteriormente, a partir do método analítico-sintético de Ranganathan, incorpora o princípio de categorização como uma solução para a organização dos conceitos num sistema de conceitos, independente da finalidade de aplicação e propõe princípios para o estabelecimento de relacionamentos associativos. Seu método leva à elaboração da definição analítica ou conceitual.

Para Motta (1987), a análise conceitual ou identificação das características facilita seu entendimento em geral, explica as relações entre conceitos em um sistema, possibilita estabelecer relacionamento e avaliar sua natureza. Portanto, a Teoria do Conceito pode ser utilizada como instrumento de auxílio em áreas onde e

quando ocorrerem problemas de clarificação e de sistematização de conceitos e pode ser aplicada a todos os casos que lidam com fundamentos de nosso conhecimento.

Medeiros (2010) afirma que, a partir da Teoria do Conceito, nota-se as relações entre os conceitos. Estas são responsáveis pela ligação entre conceitos, permitindo, por exemplo, que determinadas máquinas possam realizar inferências a partir de conceitos distintos e das relações permitidas entre eles, como é o caso das ontologias.

Medeiros e Campos (2006) apontam semelhanças na construção de tesouros conceituais (baseados na Teoria do Conceito) e ontologias quando afirmam que a Teoria do Conceito admite a existência de conceitos propriamente ditos, compostos pelo referente, suas características e um nome que o designa. O referente é um objeto no mundo, sendo classificado como objeto geral ou individual. Na ontologia formal, os objetos, ou particulares, são classificados inicialmente como *endurants* (contínuos) ou *perdurants* (ocorrentes). Os *endurants* são objetos/entidades, enquanto o *perdurants* são eventos/ações. Apesar dos tesouros conceituais não possuírem tal classificação, os conceitos que os constituem são também objetos/entidades, eventos/ações, entre outras categorias de conceitos.

A relação categorial, segundo Medeiros e Campos (2006), é apresentada na construção de tesouros através da relação formal-categorial na Teoria do Conceito. Esta relação toma por base o referente escolhido, impondo-lhe um processo de categorização e permitindo, assim, que seja montada a estrutura do domínio, conferindo estabilidade e flexibilidade a esta estrutura. Na ontologia de fundamentação, a observação parte dos objetos e a partir deles são estabelecidas relações. Nesse processo, Campos (2004, p. 28) afirma que “a categoria, especificamente, é considerada uma classe de nível mais amplo, tendo como função possibilitar uma classificação geral do domínio em questão”.

Em tesouros conceituais, de acordo com Medeiros e Campos (2006), a relação hierárquica compreende as relações gênero/espécie e a relação lateral (conceitos em renque). Na ontologia de fundamentação, a relação de gênero/espécie permite organizar taxonomicamente a estrutura do domínio. A relação partitiva ou parte-todo, na construção de tesouros conceituais, compreende a ligação entre o todo e suas partes e a relação das partes entre si. A ontologia de fundamentação, por sua vez, compreende o estudo aprofundado desta relação, destinando uma área da Filosofia para estudar especificamente os relacionamentos entre o todo e as partes de uma entidade.

A relação funcional-sintagmática é conceitualmente orientada a processos ou operações (DAHLBERG, 1978b). A ontologia de fundamentação tem, na ontologia

formal, subsídios para que esta seja trabalhada através da noção de dependência a ligação entre os conceitos, explicitando a “dependência existencial, envolvendo indivíduos específicos pertencentes a classes diferentes” (CAMPOS, 2004, p. 30). Assim, embora não nomeada desta forma, a ontologia de fundamentação possui uma vasta tipologia destas relações (MEDEIROS e CAMPOS, 2006).

Moreira (2003) sugere, em vista da ausência de uma metodologia amplamente adotada para a elaboração de ontologias, a adaptação das metodologias utilizadas na CI, particularmente a Teoria do Conceito, para ser usada como base para a construção de ontologias.

Marcondes e Campos (2008) verificam quatro princípios fundamentais que devem auxiliar o modelizador na elaboração de estruturas conceituais em domínios de conhecimento, como é o caso das ontologias, quais sejam: o método de raciocínio; o objeto de representação; as relações entre os objetos e as formas de representação. Os autores apontam nas três primeiras a possibilidade da contribuição da Teoria do Conceito.

Silva (2008) propõe uma metodologia para construção de ontologias que utilize princípios da Teoria do Conceito na análise de domínio na identificação de termos e definições de conceitos, relações e atributos. Segundo a autora, a abordagem dedutiva e indutiva da Teoria do Conceito vai ao encontro da realidade de construção de ontologias que se utilizam de métodos *top-down* para determinação de conceitos genéricos e *bottom-up* para determinação de conceitos de nível mais baixo. Silva (2008) destaca a contribuição da CI na fase de modelagem conceitual de uma ontologia no que se refere às técnicas de análise de assunto dos documentos, na análise do conceito para a realização de definições e na organização dos termos através de facetas ou categorias, viabilizando a construção da estrutura taxonômica de uma ontologia. Por fim, afirma que o aporte metodológico existente no âmbito da CI pode contribuir na proposição de metodologias para o processo de elaboração de ontologias.

5. PROPOSTA DE METODOLOGIA DE MODELAGEM CONCEITUAL PARA UMA ONTOLOGIA NO DOMÍNIO LICITAÇÕES E CONTRATOS ADMINISTRATIVOS

Metodologia é definida como “uma série integrada de técnicas ou métodos criando uma teoria geral de sistemas de como uma classe de pensamento pode ser executada”, onde método pode ser entendido como “um conjunto de processos ou procedimentos ordenados usados na engenharia de um produto ou na realização de um serviço” e técnica “um procedimento técnico e gerencial usado para alcançar um dado objetivo”. Assim, uma metodologia é composta de métodos que possuem suas próprias técnicas (FERNANDEZ, GÓMEZ-PEREZ e CORCHO, 2004 *apud* SILVA, SOUZA e ALMEIDA, 2008).

Tomando-se por base esses conceitos, pretende-se aqui propor uma metodologia, fundamentada em princípios da Teoria do Conceito para a construção de um modelo conceitual de uma ontologia no domínio¹² Licitações e Contratos Administrativos. Tal escolha baseia-se na constatação, por meio da literatura, de que dentre as metodologias de modelagem conceitual provenientes da CI esta é a mais adequada para o fim proposto.

Dentre as fases da construção da ferramenta, encontra-se a fase de modelagem conceitual onde o conhecimento sobre um dado domínio é conhecido e modelado de acordo com uma realidade. Importante ressaltar que essa realidade representa um recorte satisfatório para um determinado grupo que compartilha e aceita consensualmente a conceitualização desse domínio. A conceitualização, portanto, é válida para aquela realidade específica, podendo não valer em outro contexto.

De acordo com a literatura, nas metodologias provenientes da CC, a fase de modelagem conceitual de uma ontologia varia, mas guarda alguma semelhança no que diz respeito aos seguintes propósitos: obtenção de conhecimento sobre o domínio; coleta de conceitos; definição e estruturação de termos. Observa-se as mesmas atividades na construção de tesouros.

A seguir, uma comparação da fase de modelagem conceitual nas metodologias de construção de ontologias da CC e na metodologia proposta na CI para a construção de tesouros.

¹² A ontologia pretendida é de domínio, pois tem o objetivo de descrever um vocabulário relativo a um domínio específico.

METODOLOGIA	Gruninger e Fox	Uschold e King	Kactus	Methontology	Sensus	Método 101	Manual do BITI*
FASE	Atividades	Atividades	Atividades	Atividades	Atividades	Atividades	Atividades
MODELAGEM CONCEITUAL	Especificação de questões de competência, que objetivam especificar em linguagem natural os requisitos que a ontologia deverá ser capaz de atender; Concepção da terminologia, em que mediante declarações em lógica de primeira ordem, os conceitos e suas propriedades são organizados em uma taxonomia.	Captura ou concepção da conceitualização de ontologia.	Identificação de termos relevantes para o domínio da aplicação a partir de tais requisitos, construindo assim, um modelo preliminar; Refinamento e estruturação da ontologia a fim de obter um modelo definitivo.	Avaliação correta do modelo conceitual para evitar futuros erros no ciclo de vida da ontologia.	Identificação de termos-chave do domínio; Ligação manual dos termos-chave da ontologia; Inclusão de caminhos até o conceito de hierarquia superior; Acréscimo de novos termos para o domínio; Acréscimo de subárvores completas.	Definição de classes na ontologia; Organização das classes em uma taxonomia; Definição de propriedades para as classes e descrição de seus valores permitidos (facetas); Inclusão de valores de slots para instâncias.	Levantamento do vocabulário, que consiste em selecionar os termos representativos do assunto e defini-los de acordo com a natureza do assunto; Organização dos conceitos, que implica agrupar os termos de mesma natureza em categorias ou facetas, a fim de permitir maior compreensão do conceito e da organização das relações entre os conceitos.

*Manual do site de Biblioteconomia, Informação e Tecnologia da Informação (BITI).

Quadro 1: Comparação entre metodologias para modelagem conceitual nas Ciências da Informação e da Computação. Quadro adaptado de: SILVA, Daniela Lucas da. Um proposta metodológica para construção de ontologias: Uma perspectiva interdisciplinar entre as Ciências da Informação e da Computação. Belo Horizonte: UFMG, 2008.

5.1 Metodologia proposta

A metodologia proposta procura expor em linhas gerais as fases necessárias ao desenvolvimento de um modelo conceitual consistente para uma ontologia de domínio. Não é pretensão aqui apresentar um protótipo ou detalhar as etapas desse processo, com exemplificações e modelos, que se entende serem objeto de futuros projetos. As fases propostas no âmbito desta metodologia são:

- a) definição do escopo da ontologia;
- b) aquisição de conhecimento;
- c) identificação de termos;
- d) definição dos conceitos;
- e) agrupamento dos termos em categorias;
- f) construção da estrutura taxonômica;
- g) definição das relações;
- h) definição das instâncias.

A base conceitual da presente metodologia já foi apresentada no tópico 6.8. Portanto, o trabalho a partir deste momento buscará apresentá-la sob o aspecto mais prático, sempre com referências ao arcabouço teórico já relacionado.

5.1.1 Definição do escopo da ontologia

O processo de construção de uma ontologia toma como referência o contexto e escopo do domínio. Essa atividade requer o envolvimento dos *stakeholders*, atores envolvidos no processo, com destaque para os especialistas no domínio, e utilização de técnicas como *brainstorming* entre os envolvidos para que se extraiam os requisitos referentes à futura ontologia.

A ausência de uma delimitação precisa do escopo poderá comprometer as fases subsequentes. São imprescindíveis para o sucesso desta fase os seguintes requisitos: definição das necessidades do usuário; consulta aos profissionais corretos; busca de patrocínio junto aos tomadores de decisão para garantir a disponibilidade dos recursos; e envolvimento dos formadores de opinião.

5.1.2 Aquisição de conhecimento

Consiste no levantamento de possíveis termos para a ontologia, por meio de identificação de conceitos em linguagem natural. As fontes de informação para coleta incluem: documentos (livros, artigos, manuais), índices, gráficos, mapas conceituais, lista de termos, relatórios técnicos, teses, dissertações, legislação, jurisprudência, fontes terminológicas (tesauros, terminologias, glossários, dicionários especializados), ontologias existentes e especialistas no domínio. Como técnicas, pode-se utilizar: *brainstorming* nos contextos de indexação e uso da terminologia, mecanismos para extração automática de termos e análise de assuntos de documentos.

5.1.3 Identificação de termos

A partir dos termos coletados, deve-se criar uma lista abrangente de termos, ordená-la alfabeticamente e, determinar, mesmo que provisoriamente, a função do termo escolhido (classe, subclasse, propriedade ou relação). Essa tarefa deve ser executada com auxílio da Teoria do Conceito¹³, identificando-se termos representativos a partir do referente, que se classifica em objetos percebidos, que representarão as classes e subclasses (ex.: licitante, pregão, edital), objetos abstratos que serão as relações ou propriedades (ex.: homologa, habilita, adjudica) e objetos individuais ou instâncias de um conceito (ex.: CNPJ do licitante, número do contrato).

De início, todos os termos levantados devem ser registrados. No processo de análise dos conceitos, alguns serão descartados, por não serem representativos do

¹³ Ver p. 46.

domínio e outros serão incluídos. O que se vai organizar são os conceitos que são designados por termos. Não há termo simples ou composto, o que se considera é o conceito. O processo de definir novas classes e obter relações é iterativo e pode se repetir muitas vezes, até a obtenção de um modelo aceitável da ontologia.

5.1.4 Definição dos conceitos

A definição conceitual fornece as características do conceito, permite seu agrupamento e indica suas relações. Deve-se produzir descrições textuais precisas e sem ambiguidades para os conceitos, relações e atributos.

Dahlberg elenca como princípios para a definição: a) identificação do objeto; b) análise das características intrínsecas e extrínsecas do objeto, para definir os conceitos e as relações entre estes; c) construção da taxonomia a partir do princípio da contextualização, onde as definições e seus posicionamentos estão diretamente ligados ao domínio; d) seleção de termos que expressem os conceitos.

Deve-se considerar a definição genérica, que determina o gênero próximo e a diferença específica do conceito; a definição partitiva, que define as partes do conceito definido e a definição por extensão que lista os elementos de um conceito através de exemplos de seu uso.

Essa fase deve ser conduzida pelos especialistas no domínio, que serão capazes de traduzir o conjunto de conceitos, relações e atributos em um vocabulário consensual do domínio.

No caso de sinônimos, acrônimos ou ambiguidade, um único termo deve ser indicado para uso e os demais registrados como “sinônimos”.

5.1.5 Agrupamento dos termos em categorias

A partir do método dedutivo-indutivo que analisa o conceito dentro do contexto para posteriormente posicioná-lo no interior de uma categoria, deve-se iniciar o agrupamento dos termos. O método dedutivo (de cima para baixo) identifica as categorias conceituais que serão utilizadas para agrupar os conceitos, ele parte de uma visão geral sobre o domínio em direção aos casos mais específicos, possibilitando uma melhor compreensão de todo o processo e do domínio como um todo. Tendo-se identificado as categorias, passa-se a adotar o método indutivo (de baixo para cima), encaixando os termos coletados, que representam conceitos específicos e centrais que possam ser organizados dentro das categorias de alto nível.

Os conceitos devem ser organizados em categorias que representam características. Esse agrupamento que determina a natureza do objeto diminui os erros lógicos e facilita a identificação de qual categoria deve pertencer um

determinado objeto. Aplicar, para isso, as categorias e subcategorias formais de Dahlberg¹⁴.

Ressalta-se que nem sempre todas as categorias estarão representadas.

5.1.6 Construção da estrutura taxonômica

As relações entre os conceitos - relações hierárquicas e relações partitivas – devem ser determinadas, construindo-se assim taxonomias. Serão adotadas as metodologias *bottom-up* (que determina os conceitos de nível mais baixo para depois generalizá-los), *top-down* (determina os conceitos de nível mais alto para depois especializá-los) e *middle-out* (determina os conceitos centrais para generalizá-los e especializá-los), advindas tanto de teorias da ontologia na Ciência da Computação, quanto da Teoria do Conceito de Dahlberg¹⁵.

A relação hierárquica exprime os graus de superordenação e subordinação entre os conceitos. O termo superordenado representa o conceito mais abrangente (gênero), do qual o subordinado é um tipo (espécie). A espécie possui todas as características do gênero e mais uma, que o especializa.

Relação partitiva, por sua vez, é a relação entre o conceito no todo e suas partes.

5.1.7 Definição das relações

Aqui deverão ser estabelecidas as relações entre conceitos da mesma taxonomia ou de taxonomias diferentes. O esquema facetado contribui em tal perspectiva, porque evidencia a natureza dos conceitos e suas relações.

As relações nas ontologias correspondem aos Termos Relacionados (TR) ou Termos Associados (TA) dos tesauros. Nas ontologias, a necessidade de uma linguagem formal para processamento da máquina faz com que as relações entre os conceitos sejam explicitadas em forma de verbos ou verbos preposicionados.

No caso de relações hierárquicas, como são imutáveis, são representadas por: *tem_tipo*; *é_um* e ocorrem sempre dentro da mesma categoria.

As relações partitivas são as que revelam a relação entre o todo e suas partes e se manifestam diferentemente de acordo com a natureza do conceito. Elas também ocorrem no interior de uma mesma categoria.

¹⁴ Ver p. 46.

¹⁵ Ver p. 47.

As relações associativas serão definidas de acordo com o domínio que vão representar, e, portanto, os verbos serão selecionados de acordo com o contexto. Elas podem se manifestar entre diferentes categorias. A Norma ANSI/NISO Z39.19-2005 apresenta, de forma exemplificativa, vários tipos de relações possíveis entre categorias, sendo elas:

- Causa / Efeito
- Processo / Agente
- Processo / Contra-agente
- Ação / Produto
- Ação / Propriedade
- Ação / Alvo da ação
- Conceito ou Objeto / Propriedade
- Conceito ou Objeto / Origem
- Conceito ou Objeto / Unidade de medida ou mecanismo
- Matéria-prima / Produto
- Disciplina ou campo / Objeto ou Praticante

Ainda dentro da definição das relações devem-se estabelecer as chamadas restrições, que descrevem características limitantes para as classes. Algumas restrições possíveis são: cardinalidade da relação (define quantos valores ela pode ter, por exemplo, um aluno pode retirar cinco livros na biblioteca), o tipo de valor da relação (quais valores podem ser utilizados na relação, *strings* (cadeia de caracteres), números, booleanos, etc), a faixa de valores possíveis para relação (*range*) e as relações definidas para uma classe (*domain*)¹⁶.

5.1.8 Definição das instâncias

Instâncias são valores específicos de uma classe ou conceito e são conhecidas a partir das extensões levantadas sobre um dado conceito em um domínio. Nessa fase, devem-se definir as instâncias relevantes para os conceitos, como:

¹⁶ Em uma ontologia é preciso definir quem na hierarquia, são *domínios* e quem são *ranges* (sendo que essa definição é relativa, podendo uma classe, ora ser domínio e ora ser range em arranjos diferentes). De modo simples, se pode entender como sendo um domínio, o que pode ser entendido como o sujeito da pergunta (aquele ou aquilo que executa a ação) e o range como sendo o objeto (aquele que é executado). (DZIEKANIAK, 2010)

- Nome do atributo da instância – são os atributos de instância cujos valores podem variar para cada instância de um conceito. Ex.: (CNPJ, CPF, N° de contrato).
- Tipo de valor – determina a restrição no momento do preenchimento do campo na classe. Ex: data; decimal (nº com vírgula); double e float (números reais); integer (número inteiro); literal (dígito e letra); name (nome); string (cadeia de caracteres).
- Cardinalidade máxima e mínima – especifica o número de valores que o atributo pode receber.

6. RESULTADOS

Como resultados da aplicação da metodologia proposta, apresenta-se a seguir alguns excertos da ontologia de Licitações e Contratos Administrativos. Salienta-se que o trabalho ainda é bastante recente e que até o momento foi possível apenas iniciar o levantamento dos termos, a sua categorização e estabelecer algumas relações básicas. O *software* utilizado é o Protégé¹⁷, versão 4.1.

1. Definição das categorias Entidade, Propriedade, Atividade e Dimensão (equivalentes às classes, propriedades, relações e instâncias).

ENTIDADE (princípios, objetos imateriais e objetos materiais)	Pregoeiro, Licitante, Equipe de Apoio, Autoridade, Proposta, Edital, Requisitos de habilitação
PROPRIEDADE (quantidades, qualidades e relações)	Julgar, Habilitar, Adjudicar
ATIVIDADE (operações, estados e processos)	Julgamento da proposta, Habilitação de licitante, Adjudicação
DIMENSÃO (tempo, posição e espaço)	Prazo de entrega da documentação

2. Facetamento a partir da metodologia analítico-sintética¹⁸ para organizar o conhecimento do domínio.

Ex.: Classe Licitação possui na faceta “tipo de licitação”: melhor técnica; menor preço; técnica e preço e maior lance ou oferta; e na faceta “modalidade de licitação”: concorrência; concurso; convite; leilão; pregão e tomada de preço.

3. Estrutura taxonômica:

a) relação hierárquica de superordenação e subordinação entre os conceitos. O termo superordenado representa o conceito mais abrangente (gênero), do qual o subordinado é um tipo (espécie). A subclasse é *um tipo* da classe, ou seja, herda todas as suas características e possui uma a mais.

¹⁷ Protégé é um editor de ontologia livre, de código aberto e um framework de base de conhecimento. Fonte: Protégé, Disponível em : <<http://protege.stanford.edu/>> Acesso em 10 abr 2013.

¹⁸ Ver p. 56.

Ex.: Na classe LICITAÇÃO, na faceta “modalidade de licitação”

- Concorrência pública
- Concurso
- Convite
- Leilão
- Pregão
 - Pregão eletrônico
 - Pregão presencial
- Tomada de preço

b) relação partitiva: relação entre o todo e suas partes. A subclasse não é considerada um tipo da classe e sim uma parte que compõe o todo.

Ex.: Na classe REQUISITOS DE HABILITAÇÃO

- Capacidade técnica
- Habilitação jurídica
- Qualificação econômica-financeira
- Regularidade fiscal
- Regularidade social

4. Relações associativas: uso de verbos para caracterizar a relação entre os conceitos.

Ex.: Pregoeiro JULGA Pregão
Autoridade ADJUDICA Objeto
Melhor técnica É_TIPO_DE Licitação

7. CONCLUSÃO

No atual cenário de crescimento da demanda por sistemas de informação que permitam um melhor compartilhamento e recuperação de informação e conhecimento, as ontologias vêm se destacando como ferramenta de grande potencial para atingir esse objetivo.

A intenção deste estudo foi identificar fundamentos teóricos da organização da informação e do conhecimento capazes de contribuir na construção de ontologias e apontar o mais adequado para o alcance desse fim. A partir dessa constatação foi proposta uma diretriz metodológica a ser utilizada na construção de um modelo conceitual no domínio na área de Licitações e Contratos Administrativos.

Por meio da literatura, verificou-se que a construção de SOC's, como taxonomias e tesouros, oriundas da CI são alicerçadas em teorias e metodologias consistentes como a Teoria da Classificação Facetada e a Teoria do Conceito. A primeira inova, com a apresentação do método analítico-sintético na classificação e a segunda, com o estabelecimento da estrutura do conceito, a abordagem dedutiva-indutiva, a adoção da definição e a explicitação das relações entre os conceitos.

Desse modo, a Teoria do Conceito mostrou-se como a metodologia mais apropriada para a construção de um modelo conceitual de uma ontologia. Essa base teórica é atualmente utilizada na construção do Tesouro de Controle Externo (TECON) do TCU, razão pela qual se buscou adaptar essa *expertise* para a necessidade atual. A metodologia proposta para a construção do vocabulário é semelhante à utilizada no desenvolvimento das categorias temáticas do TECON, contudo há diferenças, considerando que se trata de sistema de organização do conhecimento distinto, com peculiaridades e objetivos diversos do tesouro.

O trabalho não buscou aprofundar na definição de um *software* específico, tendo em vista a dinâmica do mercado tecnológico que lança novas soluções frequentemente. Assim, considerando que o foco deste estudo pretendeu apresentar metodologia de construção de modelo conceitual consistente, o escopo não alcançou o desenvolvimento de protótipo ou detalhamento do processo de construção do modelo conceitual com exemplos e modelos.

Ainda que a tarefa se encontre em fase inicial, a aplicação da metodologia proposta tem se mostrado efetiva, conforme exemplificado nos excertos apresentados. E apesar de o domínio considerado ser Licitações e Contratos Administrativos, a metodologia ora proposta poderá ser aplicada a qualquer domínio de conhecimento, de modo a atender a outras demandas. Acredita-se que há uma tendência que os sistemas de informação se utilizem cada vez mais de recursos que permitam embutir

“inteligência” no ambiente digital, para promover uma melhor recuperação da informação e conhecimento, além de seu reuso e compartilhamento.

8. REFERÊNCIAS

ABBAGNANO, Nicola. Dicionário de Filosofia. Martins Fontes, 1998. Disponível em: <<http://pt.scribd.com/doc/4776000/Dicionario-de-Filosofia-Nicola-Abbagnano>>

Acesso em: 9 jan 2013.

AGUIAR, Ubiratan Diniz de, FIGUEIREDO, Alexandre, FIGUEIREDO, Pedro Ângelo Sales. O papel do controle externo: TCU, TCE, TCM. Disponível em: <http://www.tcm.ce.gov.br/site/arquivos/servicos/downloads/2010/corso_control_e_social/tcm-03.pdf> Acesso em: 24 ago 2012.

ALMEIDA, Maurício B., BAX, Marcello 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. Disponível em: <<http://www.scielo.br/pdf/ci/v32n3/19019>> Acesso em: 26 set 2012.

_____, OLIVEIRA, Viviane Nogueira Pinto de, COELHO, Kátia Cardoso. Estudo exploratório sobre ontologias aplicadas a modelos de sistemas de informação: perspectivas de pesquisa em Ciência da Informação. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/10987>> Acesso em: 22 jun 2012.

_____, et al. Uma iniciativa interinstitucional para construção de ontologia sobre ciência da informação: visão geral do Projeto P.O.I.S. R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, n.19, 1º sem. 2005. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2005v10n19p53/5500>> Acesso em: 19 out 2012.

_____, et al. Relações semânticas em ontologias: estudo de caso do Blood Project. Liinc em Revista, v.6, n.2, setembro, 2010, p. 384- 410. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/view/362/250>> Acesso em: 22 jun 2012.

ALVARENGA, Lídia. Representação do conhecimento na perspectiva da Ciência da Informação em tempos e espaços digitais. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/97/5233>> Acesso em: 25 jun 2012.

ÁLVARES, Lillian. Tesouro. Disponível em: <www.alvarestech.com/lillian/Analise/Modulo3/Aula31Tesauros.pdf> Acesso em: 20 set 2012.

ARAÚJO, C. A. A. Fundamentos teóricos da classificação. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2006v11n22p117/368>> Acesso em: 16 out 2012.

BAIÃO, Fernanda Araújo et al. Pesquisa e prática em ontologias e em modelagem de processos de negócio aplicadas à Arquitetura de Informações. Disponível em: <<http://ontobra.comp.ime.br/artigos/Gr15.pdf>> Acesso em: 19 jun 2012.

BALANIUK, Remis. A Mineração de Dados como apoio ao Controle Externo. Revista do TCU, n. 117, jan/abr 2010. Disponível em: <<http://portal2.tcu.gov.br/portal/pls/portal/docs/2057628.PDF>> Acesso em: 22 ago 2012.

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

BARQUÍN, Beatriz A., González, José Antonio, PINTO, Adilson Luiz. Construção de uma ontologia para sistemas de informação empresarial para a área de Telecomunicações. Disponível em: <http://www.dgz.org.br/abr06/Art_04.htm> Acesso em: 19 jun 2012.

BARROS, L. A. Aspectos epistemológicos e perspectivas científicas da terminologia. **Ciência e Cultura**, v. 58, n. 2, p. 22-26, 2006.

BERNERS-LEE, T., HENDLER, J., LASSILA, O. "The Semantic Web". **Scientific America**, vol. 284, 2001, pp. 34-43.

BOCCATO, Vera Regina Casari, RAMALHO, Rogério Aparecido Sá, FUJITA, Mariângela Spotti Lopes. A contribuição dos tesouros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. Disponível em: <<http://ibersid.eu/ojs/index.php/ibersid/article/view/2235/1996>> Acesso em: 18 out 2012.

BORGIDA, BRACHMAN, R. J. **Conceptual modeling with description logics**. 2003.

BRÄSCHER, Marisa. **Elaboração de tesouros**. Brasília, 2010.

_____. Tesouro, Taxonomia e Ontologia: uma evolução? Disponível em: <<http://projeto.lexml.gov.br/arqs/BRASCHER.pdf>> Acesso em: 19 jun 2012.

_____, CAFÉ, Lúcia. Organização da informação ou organização do conhecimento? Disponível em: <<http://www.ancib.org.br/media/dissertacao/1835.pdf>> Acesso em: 20 dez 2013.

BRASIL. Constituição Federal de 1988. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm> Acesso em: 12 nov 2012.

_____. Lei nº 8.443, de 16 de julho de 1992. Dispõe sobre a Lei Orgânica do Tribunal de Contas da União e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8443.htm> Acesso em: 12 nov 2012.

_____. Lei nº 8.666, de 21 de junho de 1993. Regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8666cons.htm> Acesso em: 12 nov 2012.

_____. Tribunal de Contas da União. Código de ética dos servidores do TCU. Disponível em: <http://portal2.tcu.gov.br/portal/page/portal/TCU/isc/programas_educacionais/futuros_servidores/1.3%20-%20Vis%C3%A3o%20Geral%20do%20Controle%20Externo%20-%20%20-%20C%C3%B3digo%20de.pdf> Acesso em: 8 abr 2013.

_____._____. Novos passos em busca da modernização do controle externo. Disponível em: < <http://portal2.tcu.gov.br/portal/pls/portal/docs/1/2057636.PDF>> Acesso em: 5 nov 2012.

_____._____. O TCU e o Controle Externo. Disponível em:<http://portal2.tcu.gov.br/portal/page/portal/TCU/comunidades/congresso/controle_externo> Acesso em: 24 ago 2012.

_____. _____. Portaria nº 205, de 1º de junho de 2009. Institui a Rede Interna de Informações para suporte ao controle externo (RI). Disponível em:<<http://www.tcu.gov.br/Consultas/Juris/Docs/judoc/PORTN/20090805/PRT2009-205.doc>> Acesso em: 24 ago 2012.

_____._____. Resolução nº 246, de 30 de novembro de 2011. Altera o Regimento Interno do Tribunal de Contas da União, aprovado pela Resolução TCU nº 155, de 4 de dezembro de 2002. Disponível em: <http://portal2.tcu.gov.br/portal/page/portal/TCU/normativos/sobre_normativos/regimento.pdf> Acesso em: 12 nov 2012.

BRAVO, Carlos de Oliveira. Geração Automática de Ontologias para a Web Semântica. Disponível em:<http://monografias.cic.unb.br/dspace/bitstream/123456789/250/1/UnB_CIC_Dissertacao_Mestrado_Impressao_Carlos_Bravo.pdf> Acesso em: 19 jun 2012.

BUSH, V. As we may think. *Atlantic Monthly*, v. 176, p. 101 108, 1945.

_____, MENDES, Fernanda. Estudo sobre a estrutura definitiva para desenvolvimento de ontologias. *Informação & Sociedade: estudos*, João Pessoa, v. 19, n. 2, p. 71-80, maio/ago. 2009. Disponível em: <<http://periodicos.ufpb.br/ojs2/index.php/ies/article/view/3152/3026>>. Acesso em: 22 jun 2012.

CAMILO, Cássio Oliveira, SILVA, João Carlos da. Um estudo sobre a interação entre Mineração de Dados e Ontologias. Disponível em: <http://www.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_002-09.pdf> Acesso em: 17 jul 2012.

CAMPOS, Maria Luiza de Almeida. **A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional.** 2001. 198p. Tese (Doutorado em Ciência da Informação) – Departamento de Ensino e Pesquisa do IBCT – Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

_____. **Linguagem documentária: teorias que fundamentam sua elaboração.** Niterói, Rio de Janeiro: EDUFF, 2001.

_____. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, v.33, n.1, p. 22 a 32, jan./abr. 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n1/v33n1a03.pdf>> Acesso em: 19 out 2012.

_____. O papel das definições na pesquisa em ontologia. Disponível em: <http://www.scielo.br/scielo.php?pid=S1413-99362010000100013&script=sci_arttext> Acesso em: 19 jun 2012.

_____ et al. Ontologias: representando a pesquisa na área através de mapa conceitual. Disponível em: <<http://www.enancib.ppgci.ufba.br/artigos/GT2--129.pdf>> Acesso em: 19 jun 2012.

_____, CAMPOS, Linair Maria, MEDEIROS, Jackson da Silva. A representação de domínios de conhecimento e uma teoria da representação: a ontologia de fundamentação. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/10389/9287>> Acesso em: 23 nov 2012.

_____, GOMES, Hagar Espanha. Metodologia de elaboração de tesouro conceitual: a categorização como princípio norteador. *Perspect. ciênc. inf.* v.11, n.3, set./dez. 2006. Disponível em: <http://www.scielo.br/scielo.php?pid=S1413-99362006000300005&script=sci_arttext> Acesso em: 19 jun 2012.

_____._____. Taxonomia e classificação: o princípio de categorização. *DataGramaZero*, v.9, n.4, ago. 2008. Disponível em: <http://www.dgz.org.br/ago08/F_1_art.htm> Acesso em: 25 jun 2012.

_____, SOUZA, Rosali Fernandez de, CAMPOS, Maria Luiza Machado. Organização de unidades de conhecimento em hiperdocumentos: o modelo conceitual como espaço comunicacional para a realização da autoria. Disponível em: <revista.ibict.br/ciinf/index.php/ciinf/article/download/111/92> Acesso em: 20 mar 2013.

CAPURRO, Rafael. What is information science for? ; a philosophical reflection. In: Vakkari, Perti, Cronin, Blaise. **Conceptions of library and information science**. Tempere, Taylor Graham, 1991. p. 82-93.

CAVALCANTI, Cordelia R. **Indexação & tesauro: metodologia e técnicas**. Brasília: Associação de Bibliotecários do Distrito Federal, 1978. 89 p.

CINTRA, Ana Maria et al. **Para entender as linguagens documentárias**. 2 ed. São Paulo: Polis, 2002. 92 p.

COELHO, Eduardo de Mattos Pinto, BAX, Marcello Peixoto, MEIRA JR., Wagner. Suporte de ontologias aplicadas à mineração de dados por regras de associação. Disponível em: <http://ceur-ws.org/Vol-776/ontobras-most2011_paper20.pdf > Acesso em: 17 jul 2012.

COSTA, Antonio Firmino. Classificações sociais. **Leitura**. Lisboa, v. 3, n. 2, p. 65-75, out. 1997/abr. 1998.

COUGO, P. **Modelagem conceitual e projeto de banco de dados**. 3ed. Rio de Janeiro: Campus, 1997. 284p.

CURRÁS, Emilia. **Tesauros: linguagens terminológicas**. Brasília : IBICT, 1995. 286 p.

DAHLBERG, Ingetraut. Fundamentos teórico-conceituais da teoria da classificação. **Revista de Biblioteconomia de Brasília**, v. 6, n.1, jan./jun. 1978a.

_____. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v.33, n.1, p. 11-19, 2006.

_____. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v.7, n. 2, p.101-107, 1978b.

DAVIS, H. et al. Towards in integrated environment with open hypermedia systems. In: **ACM Conference on hypertext**, 1992, Milan, Italy. Proceedings... Milan, Italy, 1992.

DODEBEI, Vera Lúcia. **Tesauro: linguagem de representação da memória documentária**. Niterói: Intexto; Rio de Janeiro: Interciência, 2002.

DZIEKANIAK, Gisele Vasconcelos. Desenvolvimento de uma ontologia sobre componentes de ontologias. Disponível em: ≤ <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/download/845/680>> Acesso em: 25 set 2012.

FERNANDEZ, M., GOMEZ-PEREZ, A., JURISTO, H. Methontology: from ontological art towards ontological engineering. 1997. Disponível em: <<http://citeseer.ist.psu.edu/context/544607/0/>>. Acesso em: 20 ago. 2007.

_____; CORCHO, O. Methodologies and methods for building ontologies. In: _____. **Ontological engineering**. London: Springer, 2004. p. 107-153.

FERNEDA, Edberto. Recuperação de informação: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. Disponível em:

<<http://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/pt-br.php>> Acesso em: 22 jun 2012.

FERREIRA, Aurélio Buarque de Hollanda. **Novo dicionário Aurélio da Língua Portuguesa**. 3. ed. Editora Positivo, 2004.

FREITAS, Frederico Luiz Gonçalves de. Ontologias e a Web Semântica. Disponível em:

<http://www.inf.ufsc.br/~gauthier/EGC6006/material/Aula%203/Ontologia_Web_semantica%20Freitas.pdf> Acesso em: 12 nov 2012.

FUJITA, M. S. L. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 1, p. 1-32, 2008.

GOMES, Hagar Espanha (Coord.). **Manual de elaboração de tesouros monolíngues**. Brasília: Programa Nacional de Bibliotecas de Instituições de Ensino Superior, 1990. 78 p.

_____. Classificação, Tesouro e Terminologia: fundamentos comuns. Disponível em: <<http://www.conexaorio.com/bit/tertulia/tertulia.htm>> Acesso em: 21 set 2012.

GONÇALVES, Julia Aparecida, SOUZA, Renato Rocha. Relações e conceitos em ontologias: teorias de Farradane e Dahlberg. Disponível em: <<http://www.uff.br/ontologia/artigos/15.pdf>> Acesso em: 19 jun 2012.

GONZÁLEZ, José Antonio Moreira. Evolução ontológica das linguagens documentárias: relato de uma experiência de curso organizado conjuntamente para o DT/SIBI-USP e o PPGCI/ECAInCID: *Revista de Ciência da Informação e Documentação*, v. 2, n. 1, p. 143-164, jan./jun. 2011. Disponível em: <revistas.ffclrp.usp.br/incid/article/viewArticle/36> Acesso em: 16 out 2012.

GUARINO, N. Understanding, Building, and Using Ontologies. **International Journal of Human- Computer Studies**. v. 46, issue 2-3, fev./mar. 1997, p. 293-310. Duluth: Academic Press, Inc., 1997.

_____. Formal ontology and information systems. In: FOIS, 1998, Trento. **Proceedings...** Trento: IOS Press, 1998. p. 3-15.

_____. The Ontological Level. In: CASI, R., SMITH, B., WHITE, G. **Philosophy and the Cognitive Science**. Viena: Holder-Pichler-Tempsky, 1994.

GUIZZARDI, Giancarlo et al. Ontologias de fundamentação, modelagem conceitual e interoperabilidade semântica. Disponível em: <<http://ceur-ws.org/Vol-728/paper6.pdf>> Acesso em: 25 set 2012.

HJORLAND, B. Fundamentals of knowledge organization. **Knowledge Organization**, v. 30, n. 2, p. 87-111, 2003.

_____. What is knowledge organization? **Knowledge Organization** , v.35, n.2/3,p.86-101.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. International Standard. **Terminology work: principles and methods**. 2. ed. 2000. (ISO 704).

_____. **Terminology work: vocabulary**. Part 1 – Theory and application. 1. ed. 2000. (ISO 1087-1).

_____. ISO 2788-1986: documentation guidelines for the establishment and development of monolingual thesauri. Disponível em: <<http://alvaroirarte.org/dokuwiki/lib/exe/fetch.php?media=thesaurus-iso.pdf> >

Acesso em: 7 jan 2013.

KABILAN, V. Ontology for information systems (O4IS) designs methodology: conceptualizing, designing and representing domain ontologies. 2007. 354p. Tese (Doutorado em Ciência da Computação e Sistemas) – Escola de Tecnologia da Informação e da Comunicação, The Royal Institute of Technology. Disponível em: <<http://www.divaportal.org/kth/theses/abstract.xsql?dbid=4513>>. Acesso em: 19 ago. 2008.

KEIZER, N. F. de, ABU-HANNA, A., ZWETSLOOT-SCHONK, J. H. M. Understanding terminological systems I: terminology and typology. **Methods of information in medicine**, v. 39, p. 16–21, 2000.

LARA, Marilda Lopez Ginez de. Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. *Ciência da Informação*. Brasília, v. 33, n. 2, p. 91-96, maio/ago. 2004. Disponível em: <<http://www.ibict.br/cienciainformacao/viewaarticle.php?id=304&layout=html>>

Acesso em: 22 set 2012.

MARRONI, Gilza Núria Brandão. Identificação e delimitação de relações associativas em tesouros: um estudo de caso na área de Direito do Trabalho. Disponível em:<http://repositorio.bce.unb.br/bitstream/10482/4919/1/2006_Gilza%20N%C3%BAria%20Brand%C3%A3o%20Marroni.pdf> Acesso em: 12 mar 2013.

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

MEDEIROS, Jackson da Silva, CAMPOS, Maria Luiza de Almeida. Tesouro conceitual e ontologia de fundamentação: análise de elementos similares em seus modelos de representação de domínios. Disponível em: <http://ceur-ws.org/Vol-776/ontobras-most2011_paper27.pdf> Acesso em: 19 jun 2012.

MEIRELLES, Hely Lopes. **Direito Administrativo Brasileiro**. 17. ed. São Paulo: Revista dos Tribunais, 1990.

MOREIRA, Alexandra. Tesouros e ontologias: estudo de definições presentes na literatura das áreas das ciências da computação e da informação, utilizando-se o método analítico-sintético. Belo horizonte, 2003. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/ihls-69uqku/alexandra_moreira.pdf?sequence=1> Acesso em: 19 jun 2012.

_____, ALVARENGA, Lídia, OLIVEIRA, Alcione de Paiva. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. Disponível em: <http://www.dgz.org.br/dez04/Art_01.htm> Acesso em: 19 jun 2012.

_____. Contribuição da terminologia na modelagem de sistemas computacionais. DataGramZero - Revista de Ciência da Informação, v.6, n.5, out 2005. Disponível em: < http://www.dgz.org.br/out05/Art_01.htm > Acesso em: 19 jun 2012.

MOTTA, Dilza Fonseca da. Método Relacional como Nova Abordagem para a Construção de Tesouros. Rio de Janeiro, SENAI/DN/DPEA, 1987. Disponível em: <http://www.conexaorio.com/biti/dilza/> Acesso em: 18 dez 2012.

_____. Conceito, teoria do conceito e inter-relações conceituais. Disponível em: <<http://www.conexaorio.com/biti/dilza/cap4.pdf>> Acesso em: 18 dez 2012.

OLIVEIRA, Laura de Lira e. **O papel das taxonomias em Ontologias de domínio:** contribuição da Ciência da Informação para estruturas taxonômicas. Projeto de Tese de Doutorado para ingresso no PPGCI UFF/IBICT. Niterói, 2007.

PADILHA, Natália, BAIÃO, Fernanda, REVOREDO, Kate. Ontologias de fundamentação apoiando o alinhamento de ontologias de domínio. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/wtdsi/2012/002.pdf>> Acesso em: 25 set 2012.

POMBO, Olga. Da classificação dos seres à classificação dos saberes. **Revista da Biblioteca Nacional de Lisboa**, n. 2, p. 19 a 33, 1998.

RAMALHO, Rogério Aparecido Sá. Representação do conhecimento e ontologias: reflexões interdisciplinares. Disponível em: < http://www.academia.edu/497389/REPRESENTACAO_DO_CONHECIMENTO_E_ONTOLOGIAS_reflexoes_interdisciplinares > Acesso em: 19 jun 2012.

_____. Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação. 2006. 120 f. Disponível em: <<http://www.cpap.embrapa.br/teses/online/DST14.pdf>> Acesso em: 23 nov 2012.

RIOS, Jocelma A. Ontologias: alternativa para a representação do conhecimento explícito organizacional. Disponível em: <http://www.cinform.ufba.br/vi_anais/docs/JocelmaRiosOntologias.pdf > Acesso em: 26 set 2012.

RIVIER, Alexis. Construção de linguagens de indexação: aspectos teóricos. **R. Esc. Biblioteconomia. UFMG**, Belo Horizonte, v.21, n.1, p.56-99, jan-jun.1992.

ROCHA, C. Alexandre Amorim. O modelo de controle externo exercido pelos Tribunais de Contas e as Proposições Legislativas sobre o tema. Disponível em: <<http://www.senado.gov.br/senado/conleg/artigos/direito/Omodelodecontroleexterno.pdf>> Acesso em: 24 ago 2012.

SALES, Luana Farias. **Ontologias de domínio**: estudo das relações conceituais e sua aplicação. 2006. Niterói: IBICT-UFF, 2006.

_____, CAMPOS, Maria Luiza de Almeida, GOMES, Hagar Espanha. Ontologias de domínio: um estudo das relações conceituais. Disponível em: <<http://www.scielo.br/pdf/pci/v13n2/a06v13n2.pdf>> Acesso em: 19 jun 2012.

SALES, Rodrigo de. Suportes teóricos para pensar linguagens documentárias. Revista Digital de Biblioteconomia e Ciência da Informação, v.5, n. 1, p 95-114, jul/dez. 2007. Disponível em: < www.brapci.ufpr.br/download.php?dd0=10622> Acesso em: 23 out 2012.

_____, CAFÉ, Ligia. Semelhanças e Diferenças entre Tesouros e Ontologias. DataGramZero - Revista de Ciência da Informação, v.9 n.4 Ago 2008. Disponível em: < http://www.dgz.org.br/ago08/Art_02.htm>. Acesso em: 22 set 2012.

SANTOS, Anderson Bestteti. Descrição e recuperação de ontologias em língua portuguesa: uma adaptação do modelo Ontology Metadata Vocabulary. Disponível em: <http://tede.pucrs.br/tde_busca/arquivo.php?codarquivo=2801> Acesso em: 19 jun 2012.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas e m Ciência da Informação**, v. 1, n. 1, p. 41-62, 1996.

SCHIESSL, Marcelo, BRÄSCHER, Marisa. Ontologia: ambiguidade e precisão. Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação, v. 17, n. esp.1, p. 125-141, 2012. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17nesp1p125/22729>> Acesso em: 29 ago 2012.

SILVA, Daniela Lucas da. Uma proposta metodológica para construção de ontologias: uma perspectiva interdisciplinar entre a Ciência da Informação e a Ciência da Computação. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/ECID-7NRQZ2/dissertacao_versao_final_setembro08.pdf;jsessionid=9252D5226A052FE1867876BC68697D5D?sequence=1> Acesso em: 19 jun 2012.

_____, SOUZA, Renato Rocha, ALMEIDA, Maurício B. Ontologias e vocabulários controlados: comparação de metodologias para construção. Ciência da Informação, v.

37, n. 3, p. 60-75, set/dez. 2008. Disponível em: <
<http://www.scielo.br/pdf/ci/v37n3/v37n3a05.pdf>> Acesso em: 7 jan 2013.

SOFTWARE ENGINEERING STANDARDS COMMITTEE OF THE IEEE COMPUTER SOCIETY. IEEE Standard for Developing Software Life Cycle Processes. December 1997. Disponível em:

<http://arantxa.ii.uam.es/~sacuna/is1/normas/IEEE_Std_1074_1997.pdf> Acesso em: 7 jan 2013.

_____. IEEE Standard Glossary of Software Engineering Terminology. **IEEE Computer Society**. New York. IEEE Std 610.121990, 1990.

SOUSA, Markson Roberto Ferreira. InfoArch: uma ontologia para modelar o domínio da Arquitetura da Informação para *Web*. Liinc em Revista, v.7, n.1, março 2011, p. 264-282. Disponível em:

<<http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/413/291>> Acesso em: 23 nov 2012.

SOWA, J. F. **Knowledge representation**: logical, philosophical and computational foundations. California : Brooks/Cole Publishing Co., 2000. 594 p.

SPITERI, Louise. The classification research group and the theory of integrative levels. *The Katharine Sharp Review*, 1995. Disponível em:

<<http://web.archive.org/web/20011222083409/alexia.lis.uiuc.edu/review/summer1995/spiteri.html>> Acesso em: 7 jan 2013.

TAKEDA, Tatiana. A importância do controle externo na administração pública. Disponível em: <http://www.jurisway.org.br/v2/dhall.asp?id_dh=1022> Acesso em: 24 ago 2012.

TEIXEIRA, Livia Marangon Duffles. Conceitualização na construção de ontologias: relações semânticas no âmbito do Blood Project. Disponível em:<

http://www.bibliotecadigital.ufmg.br/dspace/bitstream/handle/1843/ECID-82AGNQ/disserta_o_entregue_livia_marangon.pdf;jsessionid=A1C5D3A10EC5408DED3D440618892212?sequence=1> Acesso em: 19 jun 2012.

TERRA, Cláudio et al. Taxonomia: elemento fundamental para a gestão do conhecimento. Disponível em:

<<http://biblioteca.terraforum.com.br/Paginas/Taxonomia-elementofundamentalparaaGC.aspx>> Acesso em: 20 set 2012.

TRISTÃO, Ana Maria Delazari, FACHIN, Gleisy Regina Bóries, ALARCOM, Orestes Estevam. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 161-171, maio/ago. 2004.

USCHOLD, M. **Building Ontologies**: towards a unified methodology. Expert Systems '96, The 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems. Cambridge, 1996.

VALENTIN, José Roberto. Convênios com ONGs: abordagem com ênfase nas fragilidades, aplicando mineração de dados. Disponível em: <<http://portal2.tcu.gov.br/portal/pls/portal/docs/1/2434032.PDF>> Acesso em: 16 nov 2012.

VICKERY, B. C. **Classificação e indexação nas ciências**. Rio de Janeiro : BNG/Brasilart, 1980. 274 p.

VITAL, Luciane Paula. Recomendações para construção de taxonomia em portais corporativos. Florianópolis, 2007. Dissertação de mestrado. Disponível em: <<http://www.cin.ufsc.br/pgcin/Vital,%20Luciane.pdf>> Acesso em: 25 jun 2012.

_____. Ontologias e Taxonomias: convergências e divergências. Disponível em: <febab.org.br/congressos/index.php/cbbd/xxiv/paper/.../289/539> Acesso em: 20 set 2012.

WAND, Y., STOREY, V.C., WEBER, R. An ontological analysis of the relationship construct in conceptual modeling. **ACM Transactions on Database Systems**, v.24, n.4, 1999. p.494-528.