

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

ANÁLISE CONCEITUAL SOBRE AS RELAÇÕES SEMÂNTICAS EM
CIÊNCIA DA INFORMAÇÃO: CONTRIBUIÇÕES PARA O
DESENVOLVIMENTO DE ONTOLOGIAS

Julia Aparecida Gonçalves Campos

BELO HORIZONTE

2009

JULIA APARECIDA GONÇALVES CAMPOS

**ANÁLISE CONCEITUAL SOBRE AS RELAÇÕES SEMÂNTICAS EM
CIÊNCIA DA INFORMAÇÃO: CONTRIBUIÇÕES PARA O
DESENVOLVIMENTO DE ONTOLOGIAS**

Dissertação apresentada ao curso de mestrado da Escola de Ciência da Informação da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de mestre em ciência da informação.

Área de concentração: Organização e uso da informação

Orientador: Prof. Dr. Renato Rocha Souza

BELO HORIZONTE

2009

C198a Campos, Julia Aparecida Gonçalves.
2009 Análise conceitual sobre as relações semânticas em Ciência da
Informação [manuscrito]: contribuições para o desenvolvimento de
ontologias / Julia Aparecida Gonçalves Campos. – 2009.
135f., enc.: il.

Orientador: Renato Rocha Souza
Dissertação (mestrado) – Universidade Federal de Minas
Gerais, Escola de Ciência da Informação.
Bibliografia: f. 124-135.

1. Ciência da Informação – Teses. 2 Relações semânticas -
Teses. 3. Ontologias - Teses. 4. Teoria do conceito. I. Título. II.
Souza, Renato Rocha III. Universidade Federal de Minas Gerais.
Escola de Ciência da Informação.

CDU 025.4



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

FOLHA DE APROVAÇÃO

"ANÁLISE CONCEITUAL SOBRE AS RELAÇÕES SEMÂNTICAS EM CIÊNCIA DA INFORMAÇÃO: CONTRIBUIÇÕES PARA O DESENVOLVIMENTO DE ONTOLOGIAS"

Julia Aparecida Gonçalves Campos

Dissertação submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais, como parte dos requisitos à obtenção do título de "**Mestre em Ciência da Informação**", Linha de Pesquisa "**Organização e Uso da Informação (OUI)**".

Dissertação aprovada em: 27 de maio de 2009.

Por:

Prof. Dr. Mauricio Barcellos Almeida - ECI/UFMG (Presidente)

Prof. Dra. Heliana Ribeiro de Mello - FALE/UFMG

Prof. Dra. Lidia Alvarenga - ECI/UFMG

Aprovada pelo Colegiado do PPGCI

Prof. Maria Aparecida Moura
Coordenadora

Versão final Aprovada por

Prof. Mauricio Barcellos Almeida
Presidente



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

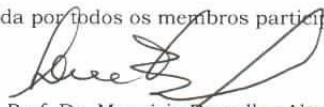
ATA DA DEFESA DE DISSERTAÇÃO DE **JULIA APARECIDA GONÇALVES CAMPOS**,
matrícula: 2007655939

Às 8:00 horas do dia 27 de maio de 2009, reuniu-se na Escola de Ciência da Informação da UFMG a Comissão Examinadora aprovada pela Câmara de Pós-Graduação em 20/05/2009, para julgar, em exame final, o trabalho intitulado **Análise conceitual sobre as relações semânticas em Ciência da Informação: contribuições para o desenvolvimento de ontologias**, requisito final para obtenção do Grau de MESTRE em CIÊNCIA DA INFORMAÇÃO, Área de Concentração: Produção, Organização e Utilização da Informação, Linha de Pesquisa: Organização e Uso da Informação. Cabe registrar que o Prof. Mauricio Barcellos Almeida assumiu a presidência da Comissão a pedido do orientador, Prof. Renato Rocha Souza (em período sabático). Esse pedido se justificou pela proximidade temática com a pesquisa da orientanda e foi devidamente aprovada por colegiado. Abrindo a sessão, o Presidente da Comissão, Prof. Dr. Mauricio Barcellos Almeida, após dar conhecimento aos presentes do teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa da candidata. Logo após, a Comissão se reuniu sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações:

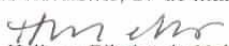
Prof. Dr. Renato Rocha Souza – orientador (por videoconferência)	APROVADA
Prof. Dr. Mauricio Barcellos Almeida - Presidente	APROVADA
Profa. Dra. Heliana Ribeiro de Mello	APROVADA
Profa. Dra. Lídia Alvarenga	APROVADA

Pelas indicações, a candidata foi considerada APROVADA.

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a sessão, da qual foi lavrada a presente ATA que será assinada por todos os membros participantes da Comissão Examinadora.



Prof. Dr. Mauricio Barcellos Almeida
Presidente - ECI/UFMG

Belo Horizonte, 27 de maio de 2009.


Profa. Dra. Heliana Ribeiro de Mello
FALE/UFMG


Profa. Dra. Lídia Alvarenga
ECI/UFMG

Obs: Este documento não terá validade sem a assinatura e carimbo da Coordenadora.


Prof.ª Maria Aparecida Moura
Coordenadora do Programa de Pós-Graduação
em Ciência da Informação - ECI/UFMG

AGRADECIMENTOS

Nada do que foi lido ou escrito, o seria, se não pela sustentação em “ombros de gigantes”, que nortearam meus ensaios científicos. Aos mestres, agradecimentos especiais para os professores: Maria Aparecida Moura; Maurício Barcellos de Almeida e Gercina Ângela Borém de Oliveira Lima. De vocês, mais que o conteúdo acadêmico, a competência e a dedicação pela área são diferenciais. Ao Prof. Renato Rocha Souza, pela orientação e estímulo à pesquisa.

Aos demais servidores da Escola de Ciência da Informação e da Escola de Educação Física, Fisioterapia e Terapia Ocupacional. Nestes espaços privilegiados, conheci pessoas com as quais construí pontes que complementaram a formação de qualidade, a qual tive acesso.

Aos amigos da graduação e pós-graduação, agradeço não apenas o apoio, mas principalmente pela amizade. O apoio de cada um foi fundamental para se cumprir mais esta tarefa; mas a amizade - presente e perene - acompanhará outros desafios. Peço licença para não citar nomes, vocês se reconhecerão nestas palavras.

À Capes, pelo apoio financeiro parcial a esta pesquisa.

Ao Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas – *Campus* Rio Pomba, pelo apoio a esta pesquisa e pelos amigos, que lá encontrei.

E seguramente, nada seria feito, senão fosse o “colo de pequenos” que, mesmo sem uma linha publicada foram primordiais para que este trabalho se tornasse realidade: meus pais, Sr. *Tarcísio* e D. *Aparecida*; meus irmãos: *Urias*, *Taciane* e *Joel*; e ao *Maurinho*, companheiro de todos os momentos, inclusive este!

Afinal, só tenho a pedir ao bom Deus – ombros e colo - que abençoe a todos. Ele sim, pode retribuir aos co-responsáveis por este trabalho.

RESUMO

Para o desenvolvimento de ontologias, verifica-se a falta de métodos sistematizados para a modelagem conceitual – resultado do processo intelectual de análise conceitual que, associado a outras técnicas, envolve a identificação, a determinação, a estruturação e a validação dos conceitos e relações semânticas apropriados para a representação da informação. Ao contrário das ferramentas tradicionais para a representação da informação, ontologias não contam com uma estrutura pré-definida e limitada de relações semânticas, mas permitem que relações semânticas refinadas sejam incorporadas à sua estrutura, de acordo com a necessidade de representação e a dinâmica do conhecimento. Por isso, análises *a posteriori* dos domínios de conhecimento a representar são essenciais para a identificação de novas relações semânticas (que aumentam de acordo com o próprio conhecimento). Ao considerar a possibilidade de identificação de uma estrutura de relações semânticas *a priori*, o objetivo deste trabalho foi servir de referência a novos projetos para o desenvolvimento de ontologias e validar a própria metodologia aplicada para a determinação de relações semânticas. Como resultado, identificou-se que não há prevalência de uma única abordagem (*a priori* ou *a posteriori*), mas há necessidade de métodos combinados para a melhor identificação da conceitualização de um domínio. O método proposto para essa verificação foi a análise conceitual, com a aplicação da teoria do conceito, onde, a partir das definições existentes na literatura para as relações semânticas determinou-se as características de cada conceito de forma criteriosa. Como resultado principal desta pesquisa identificou-se 5 grupos de relações semânticas (relações de equivalência; relações hierárquicas; relações partitivas; relações de causa-efeito e relações associativas). Em contrapartida, pesquisas desta natureza em ciência da informação corroboram o fazer metodológico da área e o aprimoramento das próprias técnicas para a organização e recuperação da informação, em resposta às necessidades atuais de representação da informação em ambiente digital.

Palavras-chave: Modelagem conceitual. Ontologias. Organização do conhecimento. Relações semânticas. Teoria do conceito.

ABSTRACT

It was verified the lack of systemized methods for the conceptual model in the ontology building process – This process, intellectual in its nature, is associated with others techniques and involves the identification, the determination, the structuring and the validation of the concepts and semantic relations appropriated for the knowledge representation. Different from the traditional tools for the knowledge representations, ontologies have no predefined and limited structure of semantic relations, but instead, allow refined semantic relations be incorporated in its structure, according to the necessity of representation and the knowledge dynamic. Because of that, a posteriori analyses of the knowledge domains to be represented are essential for the identification of new semantic relations. Considering the possibility of identification of a priori structure of semantic relations, the objective of this study was to be a reference to new ontology building projects and to validate the applied methodology to the determination of semantic relations. As a result, it was identified that there is not prevalent of unique approaches (either a priori or a posteriori), but there is the necessity of combined methods to a better identification of the dominion conceptualization. The proposed method in the methodology was the conceptual analysis with the application of the Dahlberg's Concept Theory, where, from the existing definitions in the literature for the semantic relations, it was possible to determine the characteristics of each concept in a discerning way. As a principal result in this research, it was identified five groups of semantic relations (equivalence relations; hierarchy relations; partitive relations; cause-effect relations and associative relations. In counterpart, for information science, researches of this nature corroborate for the methodological production in this area and for the improvement of the techniques for the organization and recovery of the information, in response of the real necessities of the information representation in digital ambient..

Key-words: Conceptual modeling. Ontologies. Knowledge organization. Semantic relations. Conceptual theory.

LISTA DE ILUSTRAÇÕES

Figura 01 - Relações em IDEF5 – <i>Integrated Definition Methods</i>	24
Figura 02 - Adaptado de modelo conceitual para combinar tesauros e ontologias.	31
Figura 03 – Adição das instâncias “Aspirina” e “Dor de cabeça”.	32
Figura 04- O papel axiomático das relações semânticas.	39
Figura 05 - Classificação para ontologias de Guarino (1998).	40
Figura 06 - Sistemas para a organização do conhecimento (Taxonomia de SOC)	46
Figura 07 - Exemplo da aplicação da indexação relacional de Farradane.	59
Figura 08 - Classificação das relações segundo a TGT.	62
Figura 09 - Equivalência entre as relações conceituais.	63
Figura 10 - Triângulo do conceito	79
Gráfico 01 -Classificação para ontologia de Guarino e Gangemi [2005?]	41
Quadro 01 - Ontologia hipotética baseada no Tesauro Eric	30
Quadro 02- Linguagem OWL	36
Quadro 03 - Operadores do PRECIS	52
Quadro 04 - Categorias relacionais de Farradane	56
Quadro 05 – Extrato do tesauro de economia	68
Quadro 06 - Eixos da teoria do conceito para o estudo das relações semânticas	80
Quadro 07 – Síntese de relações semânticas	84
Quadro 08 - Síntese das relações semânticas por autor de origem	86
Quadro 09 - Relação Semântica de Equivalência	109
Quadro 10 - Relação Semântica Hierárquica	109
Quadro 12 - Relação Semântica Partitiva	110
Quadro 11 - Relação Semântica Causa-Efeito	112
Quadro 12 - Relação Semântica Associativa	113

LISTA DE SIGLAS

ANSI	- <i>American National Standards Organization</i>
CC	- Ciência da computação
CDD	- Classificação Decimal de Dewey
CDU	- Classificação Decimal Universal
CI	- Ciência da informação
IBICT	- Instituto Brasileiro de Ciência e Tecnologia
IDEF	- <i>Integrated Definition Methods</i>
ISO	- <i>International Standard Organization</i>
KOS	- <i>Knowledge Organization Systems</i>
OC	- Organização do conhecimento
OI	- Organização da informação
Precis	- <i>Preserved Context Indexing System</i>
RC	- Representação do conhecimento
RDF	- <i>Resource Description Framework</i>
RDFS	- <i>Resource Description Framework Schema</i>
RI	- Representação da informação
RT	- <i>Related Term</i>
SI	- Sistema de informação
SOC	- Sistemas de organização do conhecimento
SRI	- Sistemas de recuperação de informação
Syntol	- <i>Syntagmatic Oriented Language</i>
TC	- Teoria do conceito
TE	- Termo específico
TEP	- Termo específico partitivo
TG	- Termo genérico
TGP	- Termo genérico partitivo
TGT	- Teoria geral da terminologia
TO	- Termo oposto
TOVE	- <i>Toronto Virtual Enterprise</i>
TR	- Termo relacionado

SUMÁRIO

1 INTRODUÇÃO	7
1.1 Justificativa	9
1.2 O problema da pesquisa	11
1.3 Objetivos da Pesquisa	13
1.3.1 Objetivo geral	14
1.3.2 Objetivos específicos	14
2 REFERENCIAL TEÓRICO	15
2.1 Ontologias	15
2.1.1 Metodologias para a construção de ontologias	18
2.1.2 Linguagens de representação de ontologias	34
2.1.3 Componentes e tipologia de ontologias	38
2.2 Semântica e ciência da informação	44
2.3 Relações semânticas em ciência da informação	46
2.3.1 Relações semânticas em instrumentos terminológicos	50
3 METODOLOGIA	70
3.1 Análise documental: sistematização das relações semânticas em ciência da informação	71
3.2 Análise conceitual com a aplicação da teoria do conceito: enunciados verdadeiros para as relações semânticas	77
4 RESULTADOS	82
4.1 Enunciados e características das relações semânticas	89
4.2 Considerações sobre os métodos para estabelecimento de relações semânticas em ontologias	116
5 CONSIDERAÇÕES FINAIS	121
REFERÊNCIAS	124

1 INTRODUÇÃO

O contexto digital apresenta novos desafios aos processos de produção e disseminação da informação. As formas para o acesso à informação são possíveis devido a disponibilidade online e a facilidade de publicação. Em contrapartida, há uma sobrecarga de informação acessível, não necessariamente relevante e organizada. No entanto, a singularidade do estágio atual dá-se muito mais devido às mudanças na tradicional relação, um produtor da informação e vários leitores (um para muitos), para a relação de muitos produtores para muitos leitores (muitos para muitos). Estes, os muitos leitores/autores, criam complexas redes de comunicação que, atualmente, deve nortear os processos de representação e recuperação de conteúdos neste ambiente.

Le Coadic (2004, p. 5) aponta duas características que marcam o futuro (já muito presente) da informação: sua explosão quantitativa e a implosão do tempo para a sua comunicação. A explosão quantitativa cujo crescimento é dito exponencial (o advento da escrita, quando a comunicação passou de oral para a escrita e o advento da eletrônica, que se traduziu na transição dos suportes materiais para suportes imateriais). A implosão do tempo de comunicação, visto que não há mais distância que seja obstáculo à velocidade. Esta noção parece revolucionar tanto os sistemas de informação quanto a noção precedente de explosão informacional.

A conjunção desses dois fenômenos levou ao aparecimento de fluxos [...] de informação muito elevados, isto é, a circulação de consideráveis quantidades de informação por unidade de tempo [...] (LE COADIC, 2004, p. 7).

A pesquisa em recuperação de informação, inicialmente condicionada ao acervo das bibliotecas, soma-se aos avanços do desenvolvimento tecnológico, à velocidade de processamento e aumento da capacidade de armazenamento de dados pelo computador. Ao contrário das previsões sobre a substituição de agentes humanos e supervalorização dos agentes artificiais, o trabalho executado por especialistas ganha relevância. Lancaster (2004) afirma que:

[...] até o final da década de 1940 e começo da década de 1950, a pesquisa em 'recuperação da informação' era domínio quase exclusivo da profissão de bibliotecário enquanto que, atualmente, muitos dos que trabalham nesta área parecem completamente ignorantes do fato que outros processos diferentes dos totalmente automáticos foram aplicados, com algum sucesso, à recuperação da informação durante mais de 100 anos e,

que de fato existe uma bibliografia sobre recuperação da informação além daquela da comunidade de informática (LANCASTER, 2004, p. x).

Esta pesquisa objetiva compreender os processos de organização e recuperação da informação em contexto digital, na perspectiva da ciência da informação. Há muito que se desenvolvem instrumentos terminológicos - ferramentas que propiciam a criação de representações registradas e manipuláveis a partir do conhecimento que se tem de um domínio – com a finalidade de representação do conhecimento. Acredita-se que sua base teórica pode resgatar pesquisas anteriores e adequá-las a esta nova realidade.

Apesar da relevância, os aspectos sociais não serão aprofundados nesta pesquisa, mas os assumimos como indissociáveis das questões que interessam à organização do conhecimento, pois devem refletir e atuar na própria dinâmica social, aonde a pesquisa nessa área se justifica. A análise do conteúdo dos documentos para organização e posterior recuperação avançou das primeiras pesquisas sobre coincidências lexicais para a consideração das dimensões conceituais e contextuais dos termos. Nos estudos desta natureza, ontologias aparecem como ferramentas adequadas para representação do conhecimento, de modo a garantir inteligibilidade nas comunicações não apenas entre humanos, mas também entre máquinas. Pesquisas que abordem os aspectos metodológicos e de uso destas ferramentas ganham relevância.

O presente estudo está dividido em: Introdução, no capítulo 1, são apresentadas as questões e as hipóteses que nortearam a pesquisa e exposta a necessidade da pesquisa para o estabelecimento metódico das relações semânticas em ontologias, justificado o desenvolvimento destas ferramentas como adequadas para o tratamento do conteúdo digital. No Referencial teórico (capítulo 2) fez-se a revisão de literatura necessária sobre o conceito de ontologias (componentes, tipologias e linguagens), sua inserção na web semântica e especificamente, a semântica do ponto de vista da própria disciplina semântica. Em ciência da informação apresenta-se o estudo das ferramentas para a organização do conhecimento, sob a ótica do desenvolvimento das relações semânticas. A metodologia adotada está detalhada no capítulo 3 e a sistematização dos tipos de relações; os enunciados e as características identificadas para as relações e os demais resultados encontram-se no Capítulo 4. As Considerações finais e os indicativos de estudos futuros estão descritos no Capítulo 5.

1.1 Justificativa

Para descrever o contexto digital a que nos referimos durante o desenvolvimento do texto, devemos apresentar o conceito e a filosofia da web semântica. A web semântica é um projeto em construção, em contraposição a Web Sintática. A web semântica não substitui a web Sintática, mas é, sem dúvida, um avanço da segunda. É parte do projeto gerenciado pela W3C, órgão que pretende criar mecanismos para otimizar a relação homem-máquina. Os componentes básicos para implantação da web semântica envolvem:

- a) Linguagens de marcação para a representação do conhecimento – que nada mais são que linguagens específicas para a “marcação”, ou seja, para a descrição de conteúdos digitais. Estas complementam as linguagens de apresentação anteriores, como o HTML, que fornece informação para a apresentação dos conteúdos (cores, negritos, fonte, tamanho, etc.), onde nenhum conteúdo de caráter mais semântico das páginas é descrito e passível de processamento pelo próprio computador. Portanto, são utilizadas associadas a outras linguagens como por exemplo XML, RDF, OWL entre outras;
- b) As ontologias, propriamente ditas, que ao determinar estruturas de relações entre conceitos somadas à formalização desta estrutura, geram linguagens que podem ser “compreendidas” pelos computadores através de agentes de software, e
- c) Os agentes, que são os programas criados para coletar, processar e permutar estas estruturas de relações.

A representação formal de estruturas de relações é o objetivo das ontologias, o que explica o fato de serem as mesmas apontadas como o cerne da web semântica (LEGG, 2007). Neste conjunto (linguagens de marcação + ontologias + agentes de software), as ontologias têm condições de realizar o processamento de conteúdos pelo computador, reduzindo ambigüidades semânticas. Ao fornecer uma descrição exata do conhecimento de forma diferente da linguagem natural, amplamente utilizada na web sintática onde a recuperação da informação é limitada a buscas por palavra-chave, as ontologias possibilitam o “entendimento” pelo computador. Por isso, há necessidade do desenvolvimento de ferramentas que ampliem a dimensão semântica dos termos para processamento pelo computador. As relações semânticas com propósito de representação do

conhecimento, ao serem explicitadas, eliminam interpretações dúbias: dois conceitos podem se relacionar de muitas formas diferentes, mas ao explicitar a relação pretendida, o conhecimento é comunicado.

As ontologias são modeladas basicamente a partir da identificação de um núcleo de conceitos, construídas inicialmente pela relação “é-um” para a formação de uma taxonomia. Outras relações devem ser identificadas e modeladas, de modo que a ontologia represente um domínio do conhecimento. Não é sem propósito, pois, que o eixo desta pesquisa está nos estudos sobre as relações entre os conceitos, a partir da literatura sobre o assunto em ciência da informação.

No presente trabalho, a escolha do estudo de ontologias em ciência da informação dá-se devido à formação da pesquisadora e interesse pelos estudos para a representação do conhecimento. A própria pesquisa em ontologia, por exemplo, aparece de forma descontextualizada em ciência da informação, limitada inicialmente às comparações com os tesouros. Moreira (2003), em seu trabalho de pesquisa, motivada por estas associações, faz um levantamento das definições para os dois termos na literatura e identifica a singularidade de cada instrumento, mesmo quando entendidos sob a perspectiva de instrumentos para a representação do conhecimento. Em nossa pesquisa, ao buscar trabalhos em biblioteconomia e ciência da informação, não se objetiva corroborar estas primeiras associações, mas sistematizar as contribuições da área de forma a propor melhores modelos para a representação de domínios do conhecimento.

A pesquisa em representação do conhecimento, dentro da área de organização do conhecimento, tem espaço privilegiado em ciência da informação. Reconhece-se que há um esforço da área em desenhar o seu corpo teórico, e, paralelamente às ferramentas desenvolvidas na área, os aspectos metodológicos e teóricos podem ser identificados, favorecendo a pesquisa de ferramentas atuais. No presente trabalho, os conceitos de “representação do conhecimento” e representação da informação são tratados como sinônimos, apesar dos diferentes usos e entendimentos para os mesmos, dentro da ciência da informação (já em ciência da computação, segundo Almeida (2006, p. 105), a representação do conhecimento se desenvolveu como um ramo da inteligência artificial, cuja atenção está voltada para a captura e a explicitação do conhecimento em sistemas).

Nesta perspectiva, esta pesquisa busca apresentar alguns avanços neste sentido, afinal, trata-se de “um campo dedicado às pesquisas científicas e à prática profissional voltadas à efetiva

comunicação do conhecimento entre os seres humanos”, indiferente aos suportes que se utilizam (SARACEVIC, 1996, p. 47). A comunicação do conhecimento envolve o trabalho de tratamento da informação - que é definido por Dias (2001) como

a função de descrever os documentos, do ponto de vista físico (características físicas dos documentos) e temático (descrição do conteúdo). O resultado desta atividade são as representações documentais (referências bibliográficas, resumos, termos de indexação etc) que se constituem de unidades mais fáceis de manipular [...] - se comparado ao documento em sua íntegra [...] (DIAS, 2001).

Indiferente à mudança de ambiente, contexto e ferramentas, os princípios, os problemas e a necessidade de promover a comunicação da informação permanecem inalterados. Ao aprofundar o entendimento das possíveis aplicações das ontologias, espera-se chegar a uma abordagem mais esclarecedora sobre o papel destas para a pesquisa em organização do conhecimento.

Afinal, a pesquisa sobre os relacionamentos semânticos entre os conceitos em ontologias pode resultar em benefícios para a própria ciência da informação, pois fomenta a reflexão da área de organização e representação da informação, resultando em melhorias para os processos de recuperação da informação em contexto digital. Este estudo pode subsidiar pesquisas futuras sobre o mapeamento de ontologias, pesquisas para integração de ontologias e pesquisas sobre a formalização dos relacionamentos para o processamento em ontologias. Em consequência, a área ganha com o refinamento de suas próprias técnicas e aprofundamento teórico, além de sedimentar a ciência da informação enquanto campo privilegiado para o desenvolvimento de ontologias e afirmar as mesmas como ferramentas adequadas para a representação do conhecimento.

1.2 O problema da pesquisa

O interesse pelas ontologias enquanto ferramenta de representação do conhecimento tem aumentado nos últimos anos. Seu desenvolvimento, no entanto, não se trata de uma tarefa trivial, tendo em vista que a pesquisa em ontologias carece de maior consolidação e a literatura apresenta uma dispersão de definições e de possibilidades de aplicação.

Estruturalmente, ontologias são caracterizadas por apresentar o vocabulário de um domínio formando um sistema terminológico - formado essencialmente por conceitos e por uma rica rede de relações. Adicionalmente a outros sistemas, a sua estrutura de conceitos deve ser passível de formalização e conseqüente “entendimento” pela máquina.

A pesquisa sobre o desenvolvimento metodológico das ontologias interessa ao campo da organização do conhecimento (OC) e a presente pesquisa coloca em evidência o papel das relações para o desenvolvimento das ontologias, já que ao que indica a literatura, não há uma sistematização ou um rigor metodológico para a construção de ontologias (NOY; HAFNER, 1997, KHOO; NA, 2006; BEDIN, 2007 e SILVA; SOUZA; ALMEIDA, 2006).

A OC, em ciência da informação, tem finalidade bem definida: desenvolver técnicas para a organização de registros do conhecimento com o propósito de melhorar o processo de comunicação deste conhecimento através do desenvolvimento de métodos, instrumentos, etc., que para Hjørland (2007a), são processos que se realizam por bibliotecários, especialistas em informação e também por algoritmos computacionais. Neste sentido, no presente trabalho, utilizaremos apenas o conceito de OC, por seu sentido mais geral, já que, “o objetivo do processo de organização da informação é possibilitar o acesso ao conhecimento contido na informação” (BRÄSCHER; CAFÉ, 2008, p. 5). Para as autoras:

Esses dois processos (OI e OC) produzem, conseqüentemente, dois tipos distintos de representação: a representação da informação, compreendida como o conjunto de atributos que representa determinado objeto informacional e que é obtido pelos processos de descrição física e de conteúdo, e a representação do conhecimento, que se constitui numa estrutura conceitual que representa modelos de mundo (BRÄSCHER; CAFÉ, 2008, p. 5).

O tratamento dado às relações entre os conceitos motivou a investigação da expressividade das mesmas para o desenvolvimento de ontologias robustas e, identificada a carência, esta pesquisa teve como objetivo identificar o tratamento para as relações semânticas, estudadas em ciência da informação, de forma a auxiliar o desenvolvimento dos relacionamentos para a construção de ontologias.

A fim de direcionar a pesquisa, formularam-se as seguintes questões de pesquisa:

Haveria uma estrutura essencial (nuclear) de relações entre os conceitos na construção de ontologias aplicáveis aos diferentes domínios do conhecimento?

Seria possível mapear estas relações essenciais de modo a nortear o estabelecimento das relações para a construção de ontologias?

Frente a estes questionamentos, levantamos as seguintes hipóteses:

1. A hipótese que o aporte teórico e metodológico em ciência da informação é suficientemente robusto para a pesquisa e desenvolvimento de ontologias, pelo caráter interdisciplinar e pelas pesquisas sobre o tratamento da informação.
2. Acredita-se que, a despeito da mudança de ambiente para se realizar os processos de armazenamento e busca de informações, o contexto digital não abarcou mudanças expressivas nos processos dos quais se ocupa a organização do conhecimento, ou seja, pesquisas anteriores em CI poderiam ser retomadas.
3. E a terceira hipótese compreende que, com o desenvolvimento da pesquisa em ontologias, tem-se instrumental necessário para a aplicação de estudos anteriores em biblioteconomia e CI, permitindo a efetiva aplicação dos estudos para o aprimoramento das técnicas para a organização do conhecimento.
4. Estas hipóteses são retomadas ao final desta pesquisa.

1.3 Objetivos da Pesquisa

Nesta seção, apresentam-se o objetivo geral e os objetivos específicos pretendidos neste trabalho que orientam o desenvolvimento da pesquisa.

1.3.1 Objetivo geral

Propor um núcleo essencial de relações em ontologias a partir da literatura sobre conceitos e relacionamentos da ciência da informação. A proposta é que esse núcleo fornecerá as relações significativas para auxiliar o desenvolvimento de metodologias para a representação de um domínio do conhecimento, baseadas na manipulação dos conceitos e das relações desse domínio.

1.3.2 Objetivos específicos

- Sintetizar os tipos de relações semânticas em ciência da informação para nortear seu estabelecimento para a construção de ontologias;
- Fornecer respaldo teórico para a proposição de relações em ontologias, a partir das relações caracterizadas em ciência da informação; e
- Contribuir com a sistematização dos métodos de análise e proposição de relações em ciência da informação.

2 REFERENCIAL TEÓRICO

Para a finalidade desta pesquisa, buscaram-se algumas definições para as ontologias que enfatizassem seus aspectos aplicáveis ao contexto da representação e recuperação de informação e de interesse para a pesquisa em ciência da informação: estrutura, aspectos metodológicos e possíveis aplicações das ontologias.

O levantamento e tratamento de definições auxiliam a compreensão das ontologias como ferramentas apropriadas para a organização do conhecimento e enumeram os elementos essenciais destas ontologias que melhor se apliquem à pesquisa. A partir da análise das definições teremos condições de compreender o tratamento dado às relações conceituais no desenvolvimento destas estruturas.

Este capítulo está dividido nas seguintes seções: a seção 2.1 trata o conceito de Ontologia a partir das metodologias existentes, de suas tipologia, linguagens de representação e estrutura, com ênfase na verificação do tratamento às relações semânticas. As relações semânticas, sua inserção em instrumentos terminológicos e seu papel para a representação do conhecimento em ciência da informação são tratadas na seção 2.3.

2.1 Ontologias

Inicialmente, para conceituar ontologias, preferiu-se a definição mais completa proposta por Guarino (1998, p. 2), que afirma que:

[...] uma ontologia se refere a um artefato de engenharia, que é constituído por um vocabulário específico utilizado para descrever certa realidade, mais um conjunto de suposições explícitas a respeito do significado pretendido para as palavras do vocabulário. Esse conjunto de suposições tem em geral a forma da teoria da lógica de primeira ordem, onde palavras do vocabulário aparecem com nomes de predicados unários ou binários, respectivamente chamados conceitos e relações. No caso mais simples, uma ontologia descreve uma hierarquia de conceitos relacionados genericamente; em casos mais sofisticados, axiomas são adicionados à estrutura de

forma a expressar outras relações entre conceitos, e para restringir a interpretação pretendida para tais conceitos¹ (GUARINO, 1998, p. 2, grifo nosso).

No entanto, deve-se enfatizar que a adoção do termo ontologia em ciência da informação e ciência da computação é posterior à sua origem como disciplina da Filosofia. Segundo Japiassú e Marcondes (2006, p. 206) Ontologia (do grego *onto*: o ser e *logos*: teoria) é um:

[...] termo introduzido pelo filósofo alemão Rudolph Goclenius, designando o estudo da questão mais geral da metafísica, a do ser enquanto ser, isto é, do ser considerado independentemente de suas determinações particulares e naquilo que constitui sua inteligibilidade própria. [...] O termo aparece no vocabulário filosófico por vezes como sinônimo de metafísica: ‘Os seres, tanto espirituais quanto materiais, têm propriedades gerais como a existência, a possibilidade, a duração; o exame dessas propriedades forma esse ramo da filosofia que chamamos de ontologia, ou ciência do ser ou metafísica geral’ (JAPIASSÚ; MARCONDES, 2006, p. 206).

Segundo Smith (2003), Ontologia é a ciência daquilo que existe, dos tipos e estruturas dos objetos, propriedades, eventos, processos e relações da realidade. Nas palavras do autor: “Ontologia parece providenciar uma classificação definitiva e exaustiva de entidades em todas as esferas do ser. Ontologia é um conceito utilizado por filósofos como um sinônimo de metafísica, que significa ‘além do físico’”. Para Breitman (2005, p. 30), enquanto disciplina da Filosofia, Ontologia tem por objetivo fornecer sistemas de categorização para organizar a realidade. Outros trabalhos apresentam o termo filosófico em Sistemas de Informação (ZUNIGA, 2001; SMITH *et al.*, 2003): “existem desentendimentos na literatura sobre o que é ontologia, mas ontologia em um Sistema de Informação é uma linguagem formal designada a representar um domínio particular do conhecimento e tem um propósito funcional” (ZUNIGA, 2001).

De fato, não é intenção da Filosofia a representação e formalização de suas teorias para entendimento computacional, mesmo quando a literatura aponta o uso do mesmo termo para a ciência da computação. No presente trabalho, compartilhamos a visão de autores como Guarino e Giaretta (1995) e Fonseca (2007) para os quais, enquanto teoria, ontologia em Filosofia pode

¹ No original, “an ontology refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words. This set of assumptions has usually the form of a first-order logical theory, where vocabulary words appear as unary or binary predicate names, respectively called concepts and relations. In the simplest case, an ontology describes a hierarchy of concepts related by subsumption relationships; in more sophisticated cases, suitable axioms are added in order to express other relationships between concepts and to constrain their intended interpretation” (GUARINO, 1998, p. 2).

auxiliar no desenvolvimento de ontologias de Sistemas de Informação. De acordo com Green e Rosemann (2004) se de um lado a própria introdução do termo em ciência da computação contribuiu para a fundamentação teórica nestas disciplinas, de outro não nos exime do cuidado ao referenciá-las.

Em ciência da computação, o uso do vocábulo é recente e objetiva a modelagem de sistemas interoperáveis e estruturados para o processamento de conteúdos pelo computador. A definição clássica e amplamente apresentada na literatura é de Gruber (1993, p. 199), que define ontologias como “uma especificação explícita de uma conceitualização”.

Em Vickery (1997, p. 284) temos um dos primeiros trabalhos em ciência da informação a tratar de ontologias. O autor considera as ontologias como bases de dados, com informação sobre categorias e/ou conceitos existentes no mundo/domínio, as propriedades que elas (as entidades) têm, e como se relacionam entre si. Gilchrist (2003) apresenta ontologias em trabalho de revisão, distinguindo-as de tesouros e taxonomias.

Le Coadic (2004, p. 64) apresenta o termo ontologia como um conceito técnico (da ciência da informação) que surgiu para a rede, “estabelecida sobre conhecimentos terminológicos e suas relações, conhecimento sobre um domínio, aprovados por uma comunidade comprometida em facilitar seu compartilhamento”.

Os primeiros trabalhos sobre ontologias em ciência da informação traziam forte relação entre estas e as linguagens documentárias, seja quanto à similaridade estrutural (um conjunto de conceitos e relações de um domínio), seja quanto à similaridade funcional para fins de representação e recuperação da informação (GARCIA JIMÉNEZ, 2004; QIN; PALING, 2001; MOREIRA, 2003). Ao analisar esta relação entre os instrumentos, Moreira (2003, p. 94) parte da análise das definições para o termo ontologia (em ciência da computação) e para o termo tesouro (em ciência da informação) apresentadas na literatura. A autora conclui que enquanto os tesouros focam na comunicação entre usuários e sistemas de classificação, no campo das ontologias há uma preocupação com o rigor da representação. Por isso nas ontologias privilegia-se a formalização das propriedades das relações, devido a necessidade de realizar inferências por meio do computador, enquanto nos tesouros conta-se com interventores humanos para a realização dessas inferências. Para melhor entendimento do aspecto formal em ontologias, faz-se necessário

o estudo das linguagens específicas para este fim, tratadas na seção 2.1.2, no presente trabalho, sobre linguagens de representação de ontologias.

Fonseca (2007) esclarece a relação das ontologias e os sistemas de informação, inicialmente tratadas em Guarino (1998). Para Guarino (1998, p. 10), ontologia é um dos componentes do sistema de informação que descreve um domínio genérico (ex: medicina ou automóveis) ou ferramenta (ex: diagnóstico ou vendas). Estes autores distinguem dois tipos de ontologias: ontologia *de* sistema de informação e ontologia *para* sistema de informação. A ontologia *de* sistema de informação acontece quando métodos e ferramentas emprestadas da filosofia são utilizados para ajudar a identificar as “coisas” básicas do mundo. Ontologias *para* sistemas de informação acontecem quando se dá a aplicação destas ferramentas para a descrição e organização de diferentes aspectos da realidade (no escopo de um Sistema de Informação). De acordo com Fonseca (2007), os cientistas da informação estão particularmente interessados em ontologias que descrevem sistemas de informação (ontologias *de* sistema de informação).

A partir da conceituação do termo, apresentar-se-á algumas metodologias para a compreensão do processo de construção destas ferramentas e sua conseqüente formalização.

2.1.1 Metodologias para a construção de ontologias

A escolha das metodologias não pretendeu esgotar todas as metodologias existentes, mas sim apresentar um panorama genérico de seu tratamento, limitado aos trabalhos cujas abordagens acrescentaram nossa análise. Foram considerados os seguintes trabalhos: o projeto CYC² (BREITMAN, 2005); *Enterprise Model Approach* (USCHOLD; KING, 1995); metodologia TOVE (GRUNNINGER; FOX, 1995); Projeto IDEF (PERAKATH *et al.*, 1994; FILLION *et al.*, 1995); Gene Ontology (SMITH *et al.*, 2001); Methontology (FERNANDEZ, GOMEZ-PEREZ e JURISTO, 1997); Método COMMA (GANDON, 2001) e método SENSUS (SWARTOUT *et al.*, 1997), que serão tratadas a seguir. Ademais, outros projetos para o desenvolvimento de ontologias foram considerados, de acordo com os interesses desta pesquisa (GUARINO;

² CYCORP. Disponível em: <www.cyc.com>. Acesso em: 15 jul. 2008.

WELTY, 2002; LIM, SONG; LEE, 2004; SOERGEL *et al.*, 2004; SALES, 2006; ALMEIDA, 2006; FREITAS *et al.*, 2008).

O projeto CYC (BREITMAN, 2005, p. 89) objetivou a construção de uma enorme base de conhecimento com os termos genéricos da realidade "consensual" dos humanos. Por tentar representar o conhecimento, seu nome remete a *enCYClopædia*, em inglês. O princípio básico desse sistema é obter inteligência e senso comum a partir da explicitação de uma grande quantidade de conhecimento considerado "óbvio". Esse sistema apresenta três fases distintas para a construção de ontologias:

- a) extração manual do conhecimento implícito e explícito sobre o assunto (realizada manualmente, extração a partir de fontes de conhecimento, do que está implícito no senso comum);
- b) extração do conhecimento apoiada por computadores; e
- c) extração do conhecimento gerenciada por computadores.

O modelo *Enterprise Model Approach* (USCHOLD; KING, 1995), traduzido literalmente para *Aproximação do Modelo de Empresa*, para a construção de ontologias organizacionais, foi criado a partir de uma experiência de desenvolvimento da *Enterprise Ontology* (também conhecida como *skeletal methodology*, metodologia esquelética) e fornece um método e um conjunto de ferramentas para capturar os processos e facilitar a análise. Este modelo é um exemplo de uma ontologia informal, escrita em linguagem natural.

Metodologicamente, este modelo compreende quatro atividades, descritas a seguir:

1. Identificação do propósito da ontologia;
2. Construção da ontologia:
 - a) captura: que é a identificação dos conceitos-chave e dos relacionamentos do domínio de interesse. Depois, é feita a definição textual e não ambígua dos conceitos e relacionamentos. Breitman (2005, p. 70) salienta que as descrições textuais não devem ser realizadas utilizando-se o modelo tradicional de dicionários, mas tentando maximizar a utilização de outros

termos que compõem essa descrição. Dessa forma, de modo informal, identificam-se os relacionamentos entre conceitos-chave.

b) codificação em linguagem formal: aqui os conceitos e os relacionamentos definidos na Etapa de Captura são codificados em uma linguagem formal. No entanto, ainda é necessário decidir se cada termo será modelado como classe, entidade ou relacionamento, mas nenhuma indicação para esta decisão é apresentada.

c) integração de outras ontologias, caso existam.

3. Desenvolvimento e avaliação da ontologia e;
4. Documentação: recomenda a documentação de todas as etapas realizadas para a construção das ontologias.

Para estabelecer as relações semânticas na etapa de captura do conhecimento os conceitos-chave e os relacionamentos de um domínio são identificados. Os métodos empregados para esta identificação não são esclarecidos, mas indica-se a produção de definições textuais sem ambigüidades para cada conceito e relacionamento. Para a etapa de obtenção do conhecimento, o método considera o reuso de outras ontologias. Os autores propõem a abordagem *middle-out* como estratégia para identificar os conceitos principais da ontologia, que são o uso dos processos *top-down* (processo de especificar os conceitos) e *bottom-up* (processo de generalizar os conceitos, para chegar aos conceitos mais abstratos, genéricos). Na descrição de Fernandez-Lopez (1999) a estratégia *middle-out* é apresentada como o método adequado para a identificação dos conceitos nesta metodologia, porque após a identificação dos conceitos importantes, os demais são obtidos por generalização e especialização.

Em síntese, o modelo de Uschold e King (1995) sofreu críticas por apresentar pouco suporte para o processo de conceitualização da ontologia e pela informalidade no levantamento dos termos: Breitman (2005, p. 71) salienta que há uma lacuna anterior a fase de codificação e que uma codificação intermediária deveria ser proposta, com heurísticas que auxiliassem os usuários a decidir quais termos incluir na ontologia e como classificá-los.

A metodologia TOVE - *Toronto Virtual Enterprise*, traduzido literalmente para *Empresa Virtual de Toronto*, é proposta por Gruninger e Fox (1995) e tem por proposta identificar intuitivamente os principais cenários de uma estrutura organizacional para a determinação do escopo da ontologia e chegar a um conjunto de questões em linguagem natural, as chamadas questões de competência, que orientarão a modelagem da ontologia.

O objetivo da descrição dos cenários de motivação é levantar um conjunto de problemas encontrados em uma determinada organização (sob a forma de relatos, problemas ou exemplos), que não encontram respostas nas ontologias existentes. Breitman (2005) afirma que a partir desses cenários-problemas se chega a um conjunto de soluções possíveis, que carregam a semântica informal dos objetos e relações que serão incluídos na ontologia. Na formulação informal das questões de competência faz-se a descrição das perguntas que a ontologia deve responder, baseadas nos cenários descritos.

A “especificação da terminologia” através da linguagem formal pode ser dividida em duas fases: 1) definição de um conjunto de termos/conceitos a partir de questões de competência e, 2) especificação formal dos objetos, atributos e relações da ontologia em uma linguagem formal.

As etapas de “descrição formal das questões de competência” e “especificação formal dos axiomas” utilizam linguagem formal para definir a semântica dos termos e os relacionamentos na ontologia. Na etapa de verificação da completude, *completeness* da ontologia, define-se as condições para que as soluções das Questões de Competência sejam satisfatórias.

As relações são mencionadas na etapa de “questão de competência informal” - fase de identificação do cenário através da seleção das questões sobre o contexto no qual a ontologia deve ser construída, e é feita a definição da terminologia da ontologia, ou seja, seus objetos, atributos e relações, ainda em linguagem natural) e na etapa de “especificação da terminologia em lógica de primeira ordem”, que contém a especificação de toda a terminologia definida na etapa anterior, em lógica de primeira ordem.

No projeto IDEF (*integrated definition methods*, traduzido para *métodos de definição integrados*), as ontologias desenvolvidas pretendem fornecer uma base rigorosa para o reuso e integração de modelos organizacionais (PERAKATH *et al.*, 1994; FILLION *et al.*, 1995). No

entanto, a formalização das ontologias, neste projeto não privilegia o uso da OWL, mas de uma linguagem para a formalização própria.

O projeto é composto pelos seguintes métodos: IDEFØ – modelagem de função; IDEF1 – modelagem de informação; IDEF1X – modelagem de dados; IDEF3 – captura da descrição do processo; IDEF4 – projeto orientado a objeto; IDEF5 – captura da descrição da ontologia; IDEF6 – captura racional de projeto; IDEF8 – modelagem de interface de usuário; IDEF9 – projeto de sistema de informação de cenário dirigido; IDEF10 – modelagem da arquitetura de implementação; IDEF11 – modelagem de artefato de informação; IDEF12 – modelagem de organização; IDEF13 – projeto e mapeamento do esquema três e IDEF14 – projeto de rede. Dos quais, selecionou-se o IDEFØ, IDEF1, IDEF1X, IDEF3, IDEF4, IDEF5 e IDEF6, referentes a aspectos essenciais para a modelagem conceitual³, descritos a seguir.

A modelagem de função em IDEFØ resulta em uma representação organizada das atividades e das importantes relações entre estas atividades em um modelo não-temporal e não-departamentalizado. Permite ao usuário "contar a história" do que a organização faz.

A modelagem de informação em IDEF1 é geralmente utilizada para identificar qual informação é gerida na organização e os problemas encontrados durante a análise causados pela falta de administração apropriada da informação e especificar que informação deverá ser administrada na implementação. Mais do que um método para modelagem é um método de análise utilizado para identificar:

- a) A informação coletada, estocada e gerenciada pela empresa;
- b) As regras para o gerenciamento da informação;
- c) Relacionamentos lógicos dentro da organização refletidos na informação; e
- d) Os problemas resultantes da falta de um bom gerenciamento da informação.

A modelagem de dados em IDEF1x suporta a modelagem através do uso de duas estruturas de classificação: generalização e especialização.

³ Por modelagem conceitual, entendemos todos os processos intelectuais e metodológicos necessários para a definição de um sistema de conceitos e relações para a formação de um sistema terminológico.

O método para a descrição da ontologia, o IDEF5, destina-se a apoiar a criação, a modificação e a manutenção de ontologias. Este método não sugere um "guia" para o desenvolvimento de ontologias, mas apresenta um conjunto de orientações:

1. Organização e escopo: estabelecem o propósito, a finalidade e o contexto para o desenvolvimento da ontologia.
2. Levantamento/coleta dos dados: os dados brutos necessários para o desenvolvimento de ontologia são adquiridos através do exame de documentos, mantendo-se um registro com informações sobre os documentos consultados e sua precedência.
3. Análise dos dados: a ontologia é extraída a partir dos resultados da coleta de dados. Primeiro, os objetos de interesse do domínio são listados, depois são identificados obedecendo aos limites propostos no escopo da ontologia.
4. Desenvolvimento inicial da ontologia: um projeto preliminar da ontologia é desenvolvido, o qual contém os *proto-conceitos*, isto é, descrições iniciais dos tipos, relações (*proto-relações*) e propriedades (*proto-características*). É a tentativa inicial de agrupamento dos conceitos em torno de características comuns.
5. Refinamento da ontologia e validação: os *proto-conceitos* são refinados e testados. É essencialmente um procedimento de validação dedutivo: a estrutura da ontologia é "instanciada" com dados e é comparada com a estrutura original da ontologia. Envolve o melhoramento da ontologia e finalização do processo.

As relações em IDEF5 são agrupadas em sete categorias: classificação, incluindo relações de "inclusão de classe"; meronímia; temporal; espacial; influência; dependência e relação de caso; conforme Figura 01 a seguir.

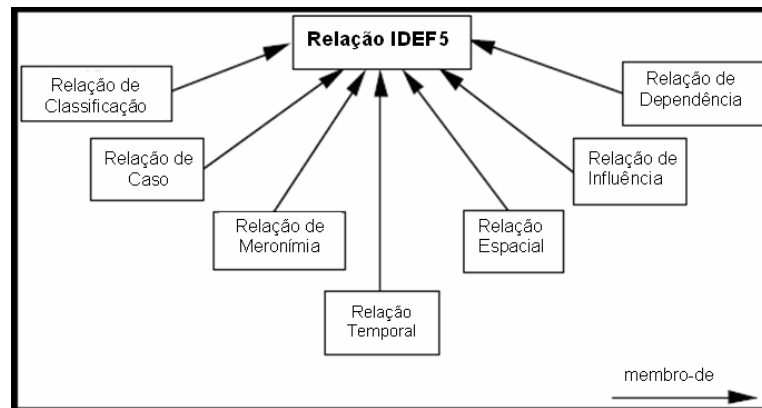


Figura 01 - Relações em IDEF5 – *Integrated Definition Methods*

Fonte: PERAKATH *et al.*, 1994.

O método apresenta algumas definições para as relações que propõe, apresentadas a seguir:

- Relação de classificação: relação de função-inclusão (martelo é uma ferramenta); relação estado-inclusão (ódio é um sentimento); relação atividade-inclusão (tênis é um esporte); relação ação-inclusão (palestrar é uma forma de falar); relação perceptual-inclusão (gato é um mamífero).
- Relação de caso: agente-ação (cachorro-latir); agente-instrumento (esquiador-esquiar); agente-objeto (padeiro-farinha); ação-receptor (deitar-cama); ação-instrumento (pintar-pincel).
- Relação de meronímia: lugar-dentro (os Alpes é parte da Europa); componente-de (roda é parte de bicicleta); coisa-de (cadeira é em parte madeira); porção-de (polegada é parte de milha); membro-de (árvore é parte de uma floresta); atividade-dentro (pagamento é parte de compra).
- Relação temporal: por ex.: antes, durante, sobrepor, encontrar, começar, e terminar.
- Relação espacial: são usadas para especificar relações espaciais entre objetos por ex.: esquerda-de, direita-de, anterior, abaixo, atrás, em-frente-de, dentro, fora, entre, longe, perto, tocante, além disso, adjacente, separar, cruzar, coincidente.

- Relação de influência: podem ser usadas para expressar que um objeto tem algum efeito ou impacto sobre outro objeto, e
- Relação de dependência: podem ser usadas para expressar o fato de um objeto depender do outro. Por ex.: depende; depende existencialmente (objeto a depende existencialmente do objeto b, se a existência de a depende da existência de b. se b cessa sua existência, então a também cessará) e depende casualmente (objeto A depende casualmente do objeto B se A é o resultado ou efeito da existência de B. Em outras palavras, B é a causa ou uma das causas para a existência de A).

Outra metodologia que se preocupa em sistematizar e normatizar as relações semânticas em ontologias é discutido em Smith *et al.* (2005) que, salientam a dificuldade em se enumerar todas as relações, o que implicaria na aceitação que algumas relações são primitivas e independentes do domínio da biologia. Nesta ontologia, esclarecem que trabalham apenas com relações ontológicas – relações que, segundo eles, podem ser verificadas na realidade.

Smith *et al.* (2005) enumeram três tipos de relações binárias⁴: a relação classe-classe, por exemplo, a relação *é_um* (*is_a*); a relação instância-classe, por exemplo, *instância_de* (*instance_of*) e a relação instância-instância, por exemplo, a relação partitiva (relação *part_of*). Assim, as relações podem ocorrer entre as classes; entre classes e instâncias e entre as instâncias propriamente. Para a modelagem conceitual do domínio, os autores se preocupam em identificar a natureza dos conceitos envolvidos na relação, de acordo com a sua materialidade e temporalidade - distinguem duas situações que envolvem alteração de uma determinada entidade, em um intervalo de tempo ou não: as entidades contínuas (*continuant entities*) e ocorrentes (*occurrent entities*).

Os termos "contínuo" e "processo" são generalizações de GO do "componente celular" e "processo biológico", mas aplicados às entidades a todos os níveis de granularidade, a partir de moléculas de todo organismo. "Contínuos" são as entidades que suportam, ou

⁴ Almeida (2006, p. 180) explica a diferença entre relações unárias (a relação entre um conceito e outro conceito, que é seu atributo, uma característica do conceito) e relações binárias (relação entre dois conceitos).

continuam a existir, através do tempo e em diferentes tipos de mudanças, incluindo mudanças de local. “Processos” são entidades que se desdobram em sucessivas fases temporais. Os termos “contínuo” e “processo”, portanto, correspondem ao que, na literatura filosófica da ontologia, são conhecidos respectivamente como "coisas" (objetos, persistente) e “ocorrentes” (atividades, eventos), respectivamente. Um “contínuo” é o que muda; um processo é a mudança em si⁵ (SMITH *et al* 2004).

Com proposta de aprimorar e reforçar uma metodologia para prover definições formais e livres de ambigüidade para as relações (e com isso, evitar erros na codificação e anotação das ontologias), existe a ontologia de relações (OBO_REL) para promover a interoperabilidade de ontologias e apoiar novos raciocínios automatizados sobre as dimensões espacial e temporal dos domínios biológico e médico.

Para Sales (2007, p. 02) a diferença entre as relações semânticas na área de ciência da informação e da ciência da computação está na forma como elas são apresentadas: o modelo de relações em ciência da informação, na esfera da elaboração de vocabulários controlados seria um modelo diádico que revelaria apenas as categorias às quais os conceitos pertencem (por ex: coisa-propriedade; material-produto; processo-resultado etc.), enquanto em ciência da computação, no que diz respeito às ontologias, o modelo apresentado seria um modelo formal que revela a forma como um elemento de uma categoria se relaciona com outro elemento de uma outra categoria. Através da análise das definições no corpus selecionado de uma ontologia da área biomédica, na Gene Ontology, a autora identificou novas relações, que não revelados em sua estrutura, estavam implícitos nestas definições. Essa análise evidenciou relações que já haviam sido sugeridas pela literatura como: *interacts_with*, *development-of*, *creates* e *part_of*, mas também encontrou relações que ainda não haviam sido mencionadas. Sales (2006) conclui que as relações são sempre um reflexo do domínio representado – ou seja, para cada ontologia a ser construída, poderão aparecer relações inéditas, ainda não mencionadas pela literatura (isto corrobora o entendimento que nem todas as relações semânticas podem ser determinadas *a priori*).

⁵ No original, “The terms 'continuant' and 'process' are generalizations of GO's 'cellular component' and 'biological process' but applied to entities at all levels of granularity, from molecule to whole organism. Continuants are those entities which endure, or continue to exist, through time while undergoing different sorts of changes, including changes of place. Processes are entities that unfold themselves in successive temporal phases. The terms 'continuant' and 'process' thus correspond to what, in the literature of philosophical ontology, are known respectively as 'things' (objects, endurants) and 'occurrents' (activities, events, perdurants) respectively. A continuant is what changes; a process is the change itself” (SMITH *et al*, 2004).

A *Methontology* (FERNANDEZ; GOMEZ-PEREZ; JURISTO, 1997) é uma metodologia desenvolvida no laboratório de Inteligência Artificial da Universidade Politécnica de Madrid (UPM), baseada da metodologia de construção de softwares. As atividades para o desenvolvimento de ontologias são divididas em três grupos:

- a) Gerenciamento de ontologias (atividades ligadas ao planejamento para a execução do projeto de construção de ontologias);
- b) Atividades ligadas ao desenvolvimento (atividades ligadas ao estudo do ambiente, viabilidade, especificação, conceitualização, formalização, integração, implementação, manutenção e uso) e
- c) Atividades de manutenção/suporte de ontologias (atividades ligadas à aquisição do conhecimento, avaliação, integração, documentação, integração, gerência da configuração e alinhamento de ontologias).

Para a análise do modo como as relações são tratadas nesta metodologia, tem-se a etapa de conceitualização e formalização. Na conceitualização é realizado o levantamento dos termos da ontologia, através de técnicas cujo objetivo é estar em contato direto com os atores que fornecem subsídios para a determinação do domínio e que têm o poder de validar o modelo e dar a conhecer a realidade a modelar. Algumas técnicas são as entrevistas, questionários, a leitura dinâmica de documentos, observação, análise de protocolo, enfoque antropológico, etc.

Na etapa de formalização, o modelo conceitual da fase anterior é formalizado através de uma linguagem, como o *OWL*, o que torna a ontologia passível de processamento automático.

O Método COMMA (*Corporate Memory Management through Agent*, traduzido para Memória associada gerenciada através de agente (GARDON, 2001) propõe uma ontologia para modelos da organização, cujas etapas de sua abordagem foram estruturadas para fins de nossa pesquisa em:

- a) Aquisição dos conceitos através de técnicas de engenharia do conhecimento. Estes conceitos e seus relacionamentos são especificados e formalizados usando *RDFS*.

- b) Faz-se uma etapa terminológica em que ocorre um controle da ambigüidade dos termos: listam-se os termos que correspondem a um único conceito; listam-se os sinônimos e os vários termos correspondendo a um conceito e identificam-se os termos que correspondem a vários conceitos (o termo é mantido, mas anotado como ambíguo). As abordagens para construção da ontologia são *bottom-up*, *top-down* e *middle-out*.
- c) A formalização da ontologia dependerá do uso pretendido para a ontologia. Utilizam-se tabelas que devem apresentar o nome do conceito/relação/atributo; as ligações de herança; os termos sinônimos; a definição por linguagem natural, na tentativa de capturar a intensão do conceito (determinada pela identificação de propriedades qualitativas e funcionais partilhadas por todas as entidades às quais o conceito se aplica).

O método SENSUS (projeto desenvolvido de 1996 a 2000) se baseou na construção de ontologias a partir de outras ontologias, identificando os termos relevantes para o domínio, ligando-os a ontologia mais abrangente. O método é baseado em abordagem *top-down* para extrair ontologias específicas (de domínios) de ontologias genéricas (*huge ontologies*, *ontologias enormes*). A proposta é a identificação de um conjunto de termos sementes (*seed*), relevantes a um domínio particular. As principais etapas são: a identificação dos termos relevantes para o domínio; a ligação dos termos considerados relevantes à ontologia *Sensus* (manualmente) e ao final, um algoritmo monta a estrutura hierárquica do domínio a partir dos termos relevantes selecionados.

Almeida (2006, p. 137) em trabalho de pesquisa para o desenvolvimento de ontologias organizacionais afirma que na primeira versão optou por inserir relações genéricas, ou seja, as relações deveriam atender a abstrações utilizadas em modelos de dados semânticos e através delas definir as relações inversas correspondentes. Essa primeira versão pareceu adequada em função de algumas constatações:

- a) primeiro, foram criadas “super-relações”(relações genéricas) para organizar as relações;

- b) segundo, obteve-se um número reduzido de relações, através de relações genéricas que buscavam similaridade com as abstrações definidas; e
- c) ao final, foi possível definir as relações inversas correspondentes (ALMEIDA, 2006, p. 180).

O trabalho de Guarino e Welty (2002) destaca o papel das relações para a constituição da espinha dorsal das ontologias. Ao explorar os erros mais comuns nas relações em ontologias, apontam para a atenção que merece a relação de *subsumption*. O termo *subsumption*, tradução subsunção, segundo o Houaiss (2001) é a ação ou efeito de subsumir, ou seja, verbo que significa incluir, colocar (alguma coisa) em algo maior, mais amplo, do qual aquela coisa seria parte ou componente. Ou ainda em filosofia:

(...) na doutrina kantiana considerar (um indivíduo) como compreendido por (uma espécie); incluir (uma espécie) em (um gênero); inserir (um gênero) em (uma família); admitir (uma idéia) como dependente de (uma idéia geral); interpretar (um fato) como a aplicação de (uma lei), e assim por diante (HOUAISS, 2001).

Guarino e Welty (2002, p. 63) apresentam a instanciação (distinção entre classes e instâncias) como um dos principais problemas encontrados nas ontologias existentes. A metodologia *Ontoclean* é tratada por análise ontológica e tem por principal benefício identificar a espinha dorsal (*backbone*) taxonômica, organizadas em relação de “subsunção” (tradução de *subsumption*; relação classe-subclasse). Através de ‘noções formais’, baseadas na ontologia filosófica, pode-se validar a estrutura taxonômica de uma ontologia, ao impor restrições na estrutura.

Soergel *et al.* (2004), ao tratar especificamente da área de agricultura, utilizam o tesouro AGROVOC (vocabulário controlado sobre agricultura, alimentos e domínios relacionados) e propõem uma “reengenharia” para que as relações existentes em tradicionais *KOS* (tesauros, esquemas de classificação, etc.) orientem o estabelecimento de relações mais refinadas para o desenvolvimento de ontologias⁶. Esta “reengenharia” implica em processos de definição da

⁶ No original, “Reengineering thesauri, classification schemes, etc., into ontologies means building on the information contained in them and refining that information as needed” (SOERGEL *et al.*, 2004).

estrutura e complementação (adição) de informação dos dados da ontologia e sua edição manual para tornar as informações mais precisas.

O quadro a seguir exhibe alguns exemplos de restrições baseadas nas relações de tesouros para uma ontologia.

Quadro 01 - Ontologia hipotética baseada no Tesouro Eric

Tesouro Eric	Ontologia Hipotética
Instrução de leitura	Instrução de leitura
BT Instrução	< é uma > instrução
RT Leitura	< tem domínio > leitura
RT Padrão de aprendizagem	< governado por > padrão de aprendizagem
Habilidade de leitura	Habilidade de leitura
BT Habilidade	< é uma > Habilidade
RT Leitura	< tem domínio > Leitura
RT Percepção	< apoiado por > Percepção

Fonte: Soergel *et al.*, 2004. Tradução nossa.

A edição manual (usando um editor de ontologia) é, segundo os autores, a etapa mais trabalhosa. Tipos especiais de restrição devem ser formulados pelo editor humano, baseado nas relações refinadas do tesouro. Baseado em padrões definidos pelo desenvolvedor, a ferramenta de edição da ontologia aplica a mesma regra imediatamente a todos os casos semelhantes.

Para a definição da estrutura, os autores apresentam um modelo (Figura 02) que pretende facilitar a transição das relações dos tesouros para ontologias. Segundo eles, sua principal característica é a clara separação do nível conceitual (designado pelo léxico) do nível do *string*, ou “cadeia de caracteres”. Compreende que um conceito pode ser representado por um ou mais termos, que pode assumir formas variáveis (singular/plural, grafia, abreviações): assim como um conceito pode ter muitas representações lexicais, um termo pode assumir várias “manifestações de *string*”⁷.

⁷ Nesta direção, a literatura apontou para Sowa (2000, p. 193) e Almeida (2006). Sowa (2000) apresenta uma forma de associação entre os vértices do triângulo semântico (OGDEN; RICHARDS, 1972) de forma a gerar, vários níveis de representação – do conceito, ao *string*, ao código – para representação dos níveis envolvidos para o processamento semântico pelo computador. Mais sobre isso, pode ser verificado em Sowa (2000) e Almeida (2006).

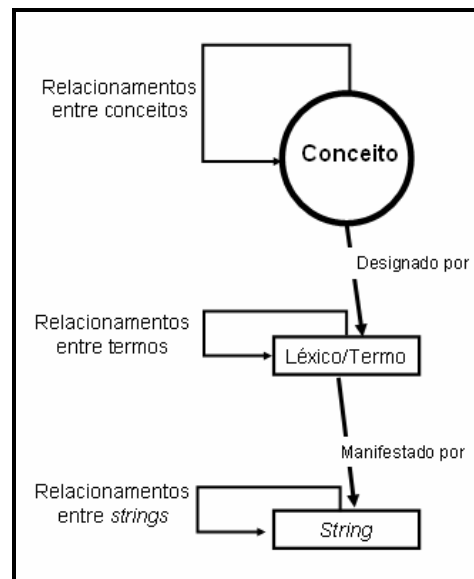


Figura 02 - Adaptado de modelo conceitual para combinar tesauros e ontologias.

Fonte: Soergel *et al.* (2004). Adaptação da autora.

Pontos importantes a considerar, segundo os autores, sobre as limitações dos *KOS* seriam: a cobertura semântica limitada e o conjunto de relações limitado, o que seriam insuficientes para guiar o usuário a localizar informação significativa na *web* ou para suportar mecanismos de inferência. Sendo as relações “imprecisas” semanticamente, poderiam ser aplicadas sem consistência, resultando em uma estrutura semântica interna irregular e imprevisível: “muitas das relações hierárquicas podem ser determinadas como relações associativas, ou vice-versa⁸ (SOERGEL *et al.*, 2004)”. No entanto, estas limitações não se mostraram como um empecilho para a determinação das relações em ontologias neste trabalho, especificamente.

Lim, Song e Lee (2004, p. 117) propõem técnicas semi-automáticas para a construção de ontologias (usando dados não-estruturados, semi-estruturados ou completamente estruturados), a partir de documentos disponíveis na *web*. Segundo eles, se um método automatizado é empregado, o núcleo da ontologia - seu sistema conceitual – deverá ser definido manualmente para garantir a qualidade na construção. No método proposto utilizam-se documentos de uma determinada área (neste caso, medicamentos), onde são extraídos os conceitos, as relações são adicionadas manualmente e expressas utilizando-se a linguagem *OWL*.

⁸ No original, “many of the NT/BT hierarchical relationships could, for example, be resolved to the non-hierarchical RT relationship, and vice versa” (SOERGEL *et al.*, 2004).

O estudo propõe dois métodos para extrair os relacionamentos entre os conceitos: extração de relações horizontais (entre objetos e agentes relevantes) usando marcações e verbos extraídos dos documentos e, extração de relações verticais (entre hiperônimos e hipônimos) entre os conceitos usando o processamento de linguagem natural. A forma como as instâncias (“Aspirina” e “DorDeCabeça”) são adicionadas à estrutura da ontologia está ilustrada na Figura 03 a seguir:

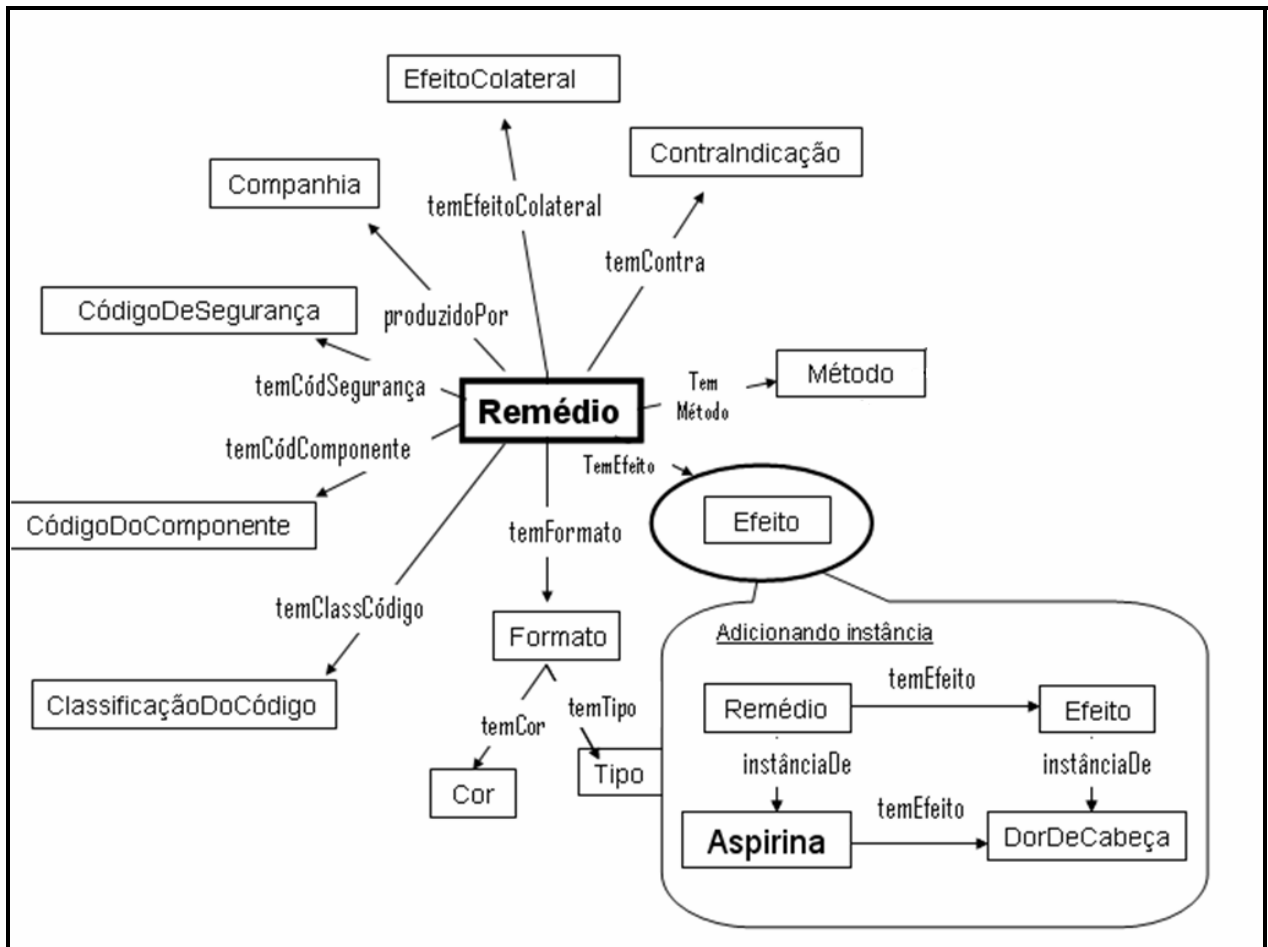


Figura 03 – Adição das instâncias “Aspirina” e “Dor de cabeça”.

Fonte: Lim; Song e Lee (2004, p. 119) Traduzido e adaptado pela autora.

A metodologia utilizada por Freitas *et al.* (2008) a partir do tesouro de Folclore e Cultura Popular Brasileira (tesouro conceitual, que agrega definições conceituais aos termos e às relações) foi possível, por apresentar definições conceituais consistentes, referendadas por um grupo de especialistas do domínio. A metodologia para a transformação do tesouro em uma ontologia ocorreu através de cinco etapas (FREITAS *et al.*, 2008, p. 03):

- a) Criação das classes: foram retirados do tesouro para a ontologia: nome, relações de equivalências e nota explicativa. Cada um dos termos presente no tesouro foi incluído na ontologia como uma Classe, utilizando-se a linguagem *OWL*.
- b) Criação das relações hierárquicas e associativas: as Relações Hierárquicas conforme modeladas no próprio tesouro. As Relações Associativas, em um primeiro momento, foram modeladas na ontologia através do uso de uma propriedade denominada “ehAssociado”, conectando-a com as classes a que ela se refere, sem aprofundar relações mais específicas, tais como “causa-efeito”, “instrumento-uso”, etc.
- c) Criação das relações partitivas: a partir da análise das definições dos termos do tesouro, as relações partitivas foram desmembradas em seis tipos: Objeto Integral/Componente; Membro/Coleção; Massa/Porção; Objeto/Matéria; Caráter/Atividade; Área/Lugar, com as respectivas relações: “Eh_Componente_de”, “Eh_Membro_de”, “Eh_Porção_de”, “Eh_Matéria_de”, “Eh_Atividade_de” e “Eh_Lugar_de”;
- d) Criação de relações obtidas por inverso de relações existentes e
- e) Organização das classes não hierarquizadas.

Para Freitas, *et al.* (2008), o ganho em se partir de ferramentas existentes está no fato de já haver uma sistematização conceitual, que resulta em ganho econômico: “a complexidade de temáticas sugerem fortemente que se deva aproveitar esforços já existentes, a fim de minimizar o custo de desenvolvimento” (FREITAS *et al.* 2008, p. 02). Além do aspecto econômico, a utilização de estruturas de conhecimento previamente organizadas sobre o domínio (em glossários, vocabulários controlados ou tesouros) pode representar um ganho significativo porque a definição textual dos termos do vocabulário fornece subsídios para compreender a natureza da associação entre eles e evidencia a semântica do domínio, além de refletir a terminologia consensual daquela área.

2.1.2 Linguagens de representação de ontologias

Para ampliar nosso entendimento sobre a formalização em ontologias e, por conseguinte, da formalização das relações entre os conceitos, parece-nos apropriado apresentar uma das principais linguagens para a formalização e determinação de axiomas em ontologias: OWL, desenvolvida especialmente para ontologias. Sem aprofundar esta questão, que exigiria conhecimento em linguagem simbólica e lógica matemática, optou-se por apresentar apenas linguagens RDF, RDFS e OWL, devido a sua inserção na já apresentada web semântica.

Reverendo o conjunto de especificações para a web semântica, tem-se que sua arquitetura é formada por camadas (BERNERS-LEE, 2001):

- a) Camada de sintaxe: XML
- b) Camada de dados: RDF
- c) Camada de ontologias: OWL
- d) Camada de lógica e camada de inferência.

De forma mais generalizada, as linguagens RDF e RDFS têm por principal característica a organização de vocabulários, hierarquicamente, através do relacionamento entre subclasses e subpropriedades, restrições de domínio e escopo e instâncias de classes. No entanto não prevê: Propriedades particulares (define uma propriedade para a classe inteira, não apenas para algumas subclasses; Disjunção de classes (por ex., as classes masculino e feminino) ou Combinação booleana de classes (união, interseção e complemento); Restrições de cardinalidade (mínimo e máximo) ou características especiais de propriedades (transitividade; exclusividade, propriedade inversa).

Por estes pontos, a expressividade do RDF e RDFS é limitada. Esforços para o desenvolvimento destas procuram fornecer uma estrutura unificada baseada em metadados para adicionar significado à web. RDF é uma forma de representar uma ontologia, enquanto que *RDF Schema* provê algumas primitivas básicas de modelagem de ontologias, como “subClassOf” (rdf:subClassOf; rdf:Class, etc) e “subPropertyOf” (rdf:Property). Outro problema em *RDF*

Schema seria a falta de semântica formal⁹ – não há mapeamento dessa linguagem para nenhum tipo de lógica formal, o que elimina qualquer possibilidade de suporte a deduções e inferências (que são fatores cruciais para processamentos automatizados).

A representação de ontologias carece de linguagens apropriadas para a formalização de sua estrutura. A OWL (*Web Ontology Language*, ou linguagem de ontologias para a Web) é a linguagem desenvolvida para representar ontologias, criada pelo Web Ontology Working Group, da W3C, mais apropriada porque se baseia na linguagem RDF e RDFS, porém define classes e relacionamentos de forma mais detalhada.

De acordo com o W3C, “a linguagem OWL é projetada para ser usada por aplicações que necessitam processar o conteúdo das informações, ao invés de apenas apresentar informações aos seres humanos” (W3C, 2004). Provê vocabulário adicional juntamente com a semântica formal, que pode ser utilizado para formalizar um domínio pela definição de classes e propriedades destas classes; definir instâncias e propriedades sobre elas e raciocinar a respeito destas classes e instâncias até o grau permitido pela semântica formal da linguagem.

Para esta pesquisa, não entraremos nas subdivisões da linguagem OWL divide suas classes em três sublinguagens, em ordem decrescente de expressividade semântica: OWL *Full*; OWL *DL* e OWL *Lite*.

- OWL *full* (tradução para completo) é a linguagem que utiliza todas as primitivas da linguagem OWL. Permite a combinação arbitrária dessas primitivas com RDF e RDFS. No entanto, fazendo este uso mais complexo não há garantia de suporte por máquinas – computabilidade;
- OWL *DL* (*Description Logic*, tradução para Lógica Descritiva) é uma sublinguagem de OWL *Full*. A vantagem é que permite suporte às inferências e a desvantagem é que perde a compatibilidade com o RDF, e

⁹ Segundo Antoniou (2004), a semântica formal suporta inferências e checka a consistência da ontologia e das relações não pretendidas entre classes. Permite raciocinar sobre o pertencimento de uma instância a uma classe; a equivalência de classes e possíveis inconsistências.

- OWL *Lite* é uma linguagem menos expressiva, com sintaxe baseada em XML e RDF, sendo, por isto, de fácil entendimento humano e fácil implementação para ferramentas.

A desvantagem, no entanto é a expressividade limitada.

Cerqueira (2007, p. 34) sistematiza os elementos da OWL para a construção de uma ontologia para o domínio do direito positivo, cujas classes, subclasses, demais relações e propriedades são sistematizadas pelo autor, conforme Quadro 02 a seguir.

Quadro 02- Linguagem OWL

Elemento	Descrição
owl:ontology	Coleção de assertivas sobre a própria ontologia, tais como suas importações, controle de versão e comentários.
owl:class	Uma classe define um grupo de indivíduos que compartilham as mesmas propriedades. A OWL define owl:Thing como sua superclasse, da qual todas são subclasses.
rdfs:subClassOf	Declara uma relação hierárquica de subclasse.
rdf:Property; owl:ObjectProperty; owl:DatatypeProperty; owl:AnnotationProperty	Propriedades (rdf:Property) estatuem relações entre indivíduos (owl:ObjectProperty) ou entre indivíduos e valores (owl:DatatypeProperty). As propriedades de anotação, que podem ser de objeto ou de dados, destinam-se basicamente à documentação da ontologia.
rdfs:subPropertyOf	Declara uma relação hierárquica de subpropriedade.
rdfs:domain	Declara uma restrição global de indivíduos na aplicação de uma propriedade.
rdfs:range	Declara uma restrição global de indivíduos que uma propriedade por ter como valor.
rdfs:label	Nome inteligível do recurso (indivíduo, classe, propriedade), destinado à leitura manual.
rdfs:comment	Descrição livre do recurso.
owl:Individual	Indivíduos são instâncias de classes.
owl:equivalentClass	Declaração de equivalência entre classes (criação de classes sinônimas).
owl:equivalentProperty	Declaração de equivalência entre propriedades (criação de propriedades sinônimas).
owl:sameAs	Declaração de identidade entre indivíduos.
owl:differentFrom	Declaração de diferença entre indivíduos.
	(Continua...)

Elemento	Descrição (Conclusão)
owl:AllDifferent	Declaração de grupo de indivíduos mutuamente distintos.
owl:inverseOf	Declaração de propriedade inversa. Dadas as propriedades P1 e P2, se P1 é declarada owl:inverseOf de P2, então $\forall x,y (P1(x,y) \leftrightarrow P2(y,x))$.
owl:TransitiveProperty	Declaração de transitividade entre propriedades. Se P é declarada owl:TransitiveProperty, então $\forall x,y,z (P(x,y) \wedge P(y,z) \rightarrow P(x,z))$.
owl:SymmetricProperty	Declaração de simetria entre propriedades. Se P é declarada owl:SymmetricProperty, então $\forall x,y (P(x,y) \leftrightarrow P(y,x))$.
owl:FunctionalProperty	Declaração de cardinalidade mínima 0 (zero) e máxima 1 (um) para a propriedade. Se P é declarada owl:FunctionalProperty, então $\forall x,y,z (P(x,y) \wedge P(x,z) \rightarrow y = z)$.
owl:InverseFunctional Property	Se P é declarada owl:InverseFunctionalProperty, então $\forall x,y,z (P(y,x) \wedge P(z,x) \rightarrow y = z)$.
owl:allValuesFrom	Quantificador universal (\forall) da Lógica Descritiva. Declaração de uma restrição local de indivíduos para uma propriedade, em relação a uma classe específica. A restrição rdfs:range é global, enquanto owl:allValuesFrom é local, alcançando apenas uma classe definida na propriedade.
owl:someValuesFrom	Quantificador existencial (\exists) da Lógica Descritiva. Declaração de uma restrição local de indivíduos para uma propriedade, em relação a uma classe específica. A restrição rdfs:range é global, enquanto owl:someValuesFrom é local, alcançando apenas uma classe definida na propriedade. Sua restrição significa que pelo menos um valor para a propriedade é do tipo especificado, sem restringir todo o escopo, como em owl:allValuesFrom.
owl:minCardinality; owl:maxCardinality	A cardinalidade diz respeito a uma restrição mínima (owl:minCardinality) ou máxima (owl:maxCardinality) de existência da propriedade nas instâncias das classes aplicáveis.
owl:cardinality	Mera propriedade de conveniência para os casos de owl:minCardinality 0 e owl:maxCardinality 0 ou owl:minCardinality 1 e owl:maxCardinality 1.
owl:unionOf; owl:complementOf; owl:intersectionOf	Operadores de conjuntos para a definição de classes complexas. Permitem a declaração de classes a partir das operações de união (\cup), negação (\neg) e interseção (\cap) de classes.
owl:oneOf	Especificação de classe pela enumeração direta de seus membros.
owl:hasValue	(property values): A property can be required to have a certain individual as a value (also sometimes referred to as property values). For example, instances of the class of <i>dutchCitizens</i> can be characterized as those people that have the Netherlands as a value of their nationality. (The nationality value, theNetherlands, is an instance of the class of Nationalities).
owl:disjointWith	Declaração de disjunção de classes.

Fonte: Conforme o original de Cerqueira (2007, p. 34-35).

2.1.3 Componentes e tipologia de ontologias

O entendimento dos tipos de ontologias é necessário para se estabelecer uma co-relação sobre os tipos de relações semânticas para as diferentes estruturas. Existem inúmeras tentativas e este trabalho buscou apresentá-las sem se ocupar em criar uma nova terminologia. No entanto, a apresentação dos tipos de ontologias prescinde do conhecimento dos componentes destas ferramentas: as classes, organizadas em uma taxonomia; os axiomas; as instâncias (os próprios dados) e as relações entre os conceitos (GRUBER, 1996), apresentados a seguir.

A parte conceitual é formada por um vocabulário, formando uma taxonomia - verificamos aí, a aplicação da Teoria da Classificação para a formação da taxonomia. Uma vez formada a taxonomia, os conceitos podem ser combinados entre si, de acordo com as relações verificadas no domínio que se pretende representar. Segundo Lim, Song e Lee¹⁰ (2004, p. 117) ontologia significa termos usados em um domínio específico, a definição de relacionamentos entre os termos e na expressão dos relacionamentos em uma estrutura hierárquica.

As instâncias são conceitos pertencentes ao domínio representado e reproduzem objetos de determinada classe. Diferem-se das classes, por serem conceitos individuais, enquanto que o pré-requisito para delimitarmos uma classe são conceitos genéricos, cuja característica é comum aos vários conceitos. Importa que, enquanto conceito individual, o conceito de herança é preservado e parte da estrutura formalizada é desenhada - ou seja, a formalização da ontologia é anterior ao uso de linguagens específicas: a modelagem de relações garante que certos comprometimentos ontológicos sejam respeitados.

O conceito de classes em ontologias é importante ao se considerar que as subclasses e instâncias herdam propriedades – segundo Campos, Souza e Campos (2005),

o conceito de herança é um dos conceitos mais poderosos no desenvolvimento de software. As máquinas podem compreender corretamente relacionamentos de generalização e especialização entre as entidades atribuindo propriedades às classes gerais e então assumindo que as subclasses herdam estas propriedades (CAMPOS, SOUZA; CAMPOS, 2005).

¹⁰No original, “ontology means terms used in a specific domain, the definition of relationships among the terms, and the expression of the relationships in a hierarchical structure” (LIM; SONG; LEE, 2004, p. 117).

Os axiomas são relacionamentos complexos, expressos em linguagem formal para o “entendimento” do computador. São assertivas expressas de acordo com as formas de relacionamentos possíveis para um determinado domínio, conforme modelos conceituais específicos, desenvolvidos por uma comunidade.

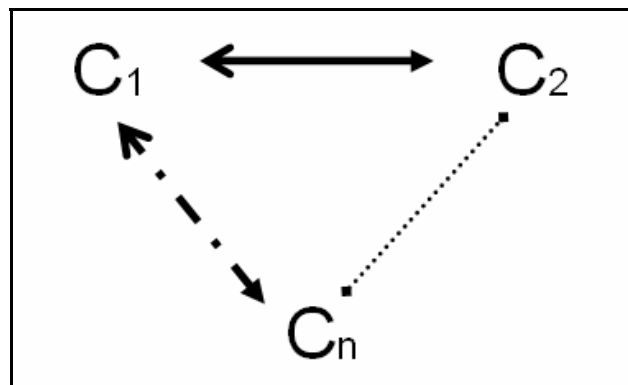


Figura 04- O papel axiomático das relações semânticas.

Na figura, as setas indicam relações: C1 relaciona-se a C2 e C1 relaciona-se a Cn. A relação, por si, já restringe o significado de um conceito na ontologia.

Fonte: da autora.

Em Oliveira (1999, *apud* ALVES, 2003, p. 9) o objetivo de axiomas é especificar a semântica de termos na ontologia e restrições na sua interpretação.

Por exemplo, em uma ontologia sobre animais africanos, pode-se especificar um axioma para o conceito de elefante adulto como sendo um elefante que possui entre cinco e oito toneladas. Nessa mesma ontologia, pode-se também dizer que o conceito vegetal é distinto do conceito animal. (OLIVEIRA, 1999, *apud* ALVES, 2003, p. 9)

Propriedades descrevem as várias características e atributos de cada conceito, e acabam por definir as características das classes. São fundamentais para a representação dos aspectos principais dos conceitos pertencentes a um domínio. Na literatura, de modo geral, não se percebe um tratamento distintivo para propriedades dos conceitos e para as relações possíveis entre os conceitos. Esta combinação vai além de relações hierárquicas: outras relações podem ser criadas.

Sobre os tipos de ontologias, Fox e Gruninger (1998, p. 112), distinguem ontologias baseadas no nível de formalidade em que seu significado é especificado: ontologias informais usam linguagem natural; ontologias semi-formais que são ontologias com poucos axiomas (como as taxonomias) e ontologias formais, que definem a semântica do vocabulário através de axiomas completos.

Guarino (1998) classifica as ontologias de acordo com sua dependência em relação a uma tarefa específica (FIG. 05), explicados a seguir.

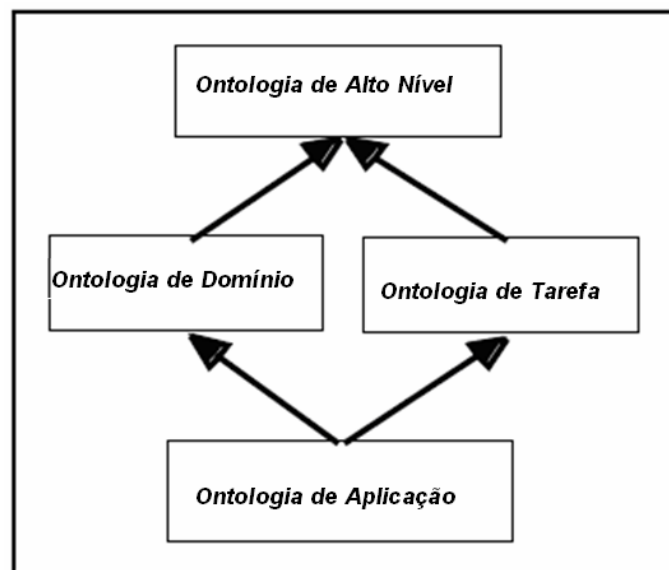


Figura 05 - Classificação para ontologias de Guarino (1998).

Fonte: GUARINO, 1998.

Ontologias de Alto Nível (*top-level ontologies ou ontologias de nível superior*): descrevem conceitos gerais e, a princípio, independem de um domínio ou tarefa, como os conceitos de espaço, tempo e eventos. São pelo menos em teoria aplicáveis a diferentes comunidades de usuários.

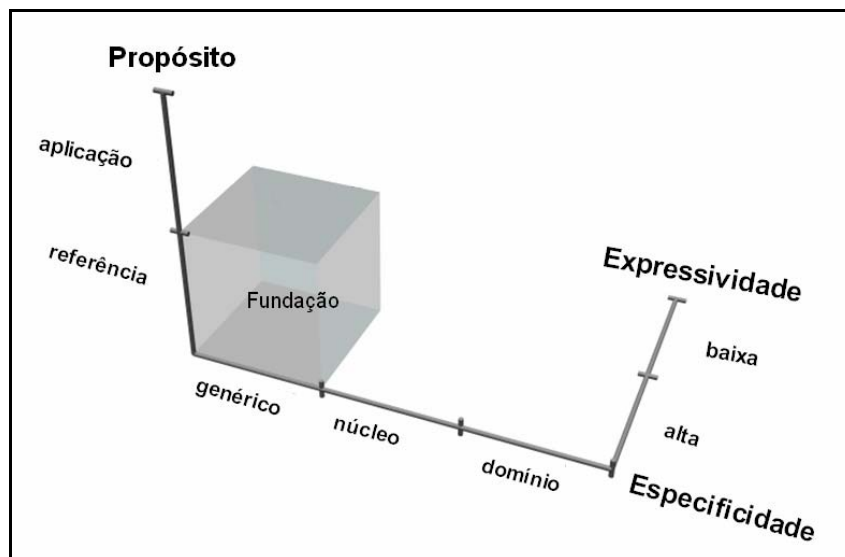
Ontologias de Domínio: descrevem um vocabulário relacionado a um domínio genérico através da especialização dos conceitos presentes na ontologia de alto nível, segundo Breitman (2005). É construída a partir da especialização dos termos introduzidos na ontologia de alto nível.

Ontologias de Tarefas: descrevem uma tarefa ou uma atividade, também através da especialização dos conceitos presentes na ontologia de alto nível, segundo Breitman (2005). É construída a partir da especialização dos termos introduzidos na ontologia de alto nível.

Ontologias de Aplicação: descrevem conceitos que dependem tanto de um domínio específico como de uma tarefa específica e é geralmente a especialização de ambos. Seus conceitos correspondem aos papéis exercidos pelas entidades de um domínio.

Recentemente, Guarino e Gangemi [2005?] apresentam o seguinte gráfico, que coloca em evidência, em ontologias, o propósito fundamentalmente referencial em contraposição à aplicação do modelo a uma tarefa específica; a generalidade dos conceitos em contraposição aos conceitos específicos de um domínio e a necessidade de um auto nível de formalidade (*heavyweight*) – o que desconsideraria os tipos *terminológicas* e *semi-formais*, como tipos válidos para ontologias.

Gráfico 1 -Classificação para ontologia de Guarino e Gangemi [2005?]



Fonte: GUARINO; GANGEMI, [2005?].

Conhecidos os tipos de ontologias, algumas das metodologias existentes e as linguagens disponíveis para o seu desenvolvimento, podemos adentrar no aspecto que norteia toda a nossa

pesquisa: as relações semânticas. Antes, porém, faz-se necessário introduzir algumas linhas de pesquisa em semântica: semântica formal, semântica da enunciação e a semântica cognitiva, apresentadas brevemente a seguir.

A semântica formal descreve o problema do significado a partir do postulado de que as sentenças se estruturam logicamente e retoma a análise de Aristóteles ao mostrar que há relações de significado que se dão independentemente do conteúdo das expressões (OLIVEIRA, 2000, p.19).

No exemplo clássico:

Todo homem é mortal.

João é homem.

Logo, João é mortal.

Esta é uma relação entre conjuntos, onde as relações se estabelecem independentemente do que homem ou mortal significam: a partir das sentenças (a) e (b), conclui-se a sentença (c), podendo ainda, serem representadas por letras vazias de conteúdo, mas que descrevem relações de sentido, assim “se A é um conjunto qualquer que está contido em um outro conjunto qualquer, o conjunto B, e se c é um elemento do conjunto A, então, c é um elemento do conjunto B” (OLIVEIRA, 2000, p. 20).

Outro aspecto importante do modelo formal está na definição de significado de Frege(1848-1925) - lógico, alemão, não aprofundado no presente trabalho, ao distinguir sentido e referência. Para esclarecer a diferença entre sentido e referência, faz-se a analogia com um telescópio apontado para a lua:

A lua é a referência: sua existência e propriedades independem daquele ou daquela que a observa, mas pode ser olhada a partir de diferentes perspectivas, e observá-la de um ângulo pode nos ensinar algo novo sobre ela. A imagem da lua formada pelas lentes do telescópio é o que tanto eu quanto você vemos. Essa imagem compartilhada é o sentido. Ao mudarmos o telescópio de posição, vemos uma face diferente da mesma lua, alcançamos o mesmo objeto por meio de outro sentido. Lembremos que a imagem mental que cada um de nós forma da imagem objetiva do telescópio está fora dos interesses da Semântica (OLIVEIRA, 2000, p. 22).

Desta forma, o sentido só nos permite conhecer algo se a ele corresponder uma referência no mundo: o sentido permite alcançarmos um objeto no mundo, mas é o objeto no mundo que nos permite formular um juízo de valor, que nos permite avaliar se o que dizemos é falso ou é verdadeiro.

Para a Semântica da enunciação, a Semântica formal é inadequada porque se respalda num modelo informacional, em que o conceito de verdade é externo à linguagem – que permitiria falar objetivamente sobre o mundo; afora da linguagem. Enquanto que, a Semântica da enunciação postula que no mundo dá-se o conteúdo para a linguagem e que esta (a linguagem) constitui o mundo e por isso, não é possível sair fora dela (OLIVEIRA, 2000, p. 27).

A Semântica cognitiva descreve a sentença a partir da hipótese de que formamos espaços mentais – a partir de nossas ações no mundo, apreendemos diretamente esquemas imagéticos espaciais e são esses esquemas que dão significado às nossas expressões lingüísticas. Esta abordagem combate a idéia de que a linguagem está numa relação de correspondência direta com o mundo, mas ao contrário, emerge de dentro para fora e por isto ele é motivado. Diferente da Semântica da enunciação, a Semântica cognitiva não se baseia na crença de que a referência é constituída pela própria linguagem nem na crença de que a linguagem é um jogo de argumentação (OLIVEIRA, 2000, p. 34). A hipótese central é de que o significado seja natural e experiencial, sustentado na constatação de que ele se constrói a partir de nossas interações físicas, corpóreas, com o meio ambiente em que vivemos.

Outro aspecto da Semântica cognitiva refere-se à questão da categorização, cara à Semântica formal (e também à ciência da informação). Pela Semântica formal, determina-se que o termo genérico *homem* não se refere a um indivíduo em particular, mas a todos os indivíduos que possam ser alcançados por meio de certas propriedades que só os humanos têm (sua intensão), que nos permite alcançar a classe de objetos no mundo (sua extensão). Segundo Oliveira (2000, p. 38), a semântica cognitiva nega a abordagem clássica de categoria. Ela se ancora fortemente em evidências psicológicas para assegurar que não categorizamos por meio do estabelecimento de propriedades necessárias e suficientes.

Como podemos adiantar, aspectos como a categorização; os silogismos de Aristóteles; modelos de significação, intensão e extensão, a abordagem sensório-cognitiva para o aprendizado das

relações em Farradane, entre outros, são conceitos comuns à pesquisa em ciência da informação – o que indica a interseção da pesquisa nas duas áreas. Apesar da aparente preferência demonstrada pela semântica formal para o desenvolvimento e formalização das ontologias, há que se reconhecerem aspectos das demais linhas da semântica, que podem contribuir para consolidar a pesquisa em ontologias. Estes apontamentos, oriundos da semântica, objetivam instigar futuras pesquisas.

2.2 Semântica e ciência da informação

Hjørland (2007c), ao ressaltar a importância da Semântica em ciência da informação, em particular para a organização do conhecimento, considera primeiramente se uma determinada relação semântica entre dois termos é constante em todos os contextos. Assim, a questão que se verifica na literatura é se, seriam as relações semânticas determinadas *a priori* (impondo-se uma ordem teórica) ou se apenas *a posteriori* (partindo-se das relações concretas entre as “coisas”).

Então, um problema fundamental na organização do conhecimento é determinar se relações semânticas podem ser determinadas *a priori* ou *a posteriori*, se podem ser conhecidas antes do exame da literatura ou somente após; que tipo de garantia literária¹¹ (ou de outro tipo) são necessárias para identificar relações e classificar conceitos¹² (HJORLAND, 2007c, p. 378)

Este questionamento é retomado de Frohmann (1983, p. 22), onde relações semânticas podem ser identificadas *a priori*, como em Austin (1974), por se basear no significado de cada termo indicado na relação - o sentido como algo permanente ou se apenas identificadas *a posteriori*, são empíricas, como defende o CRG (1955 apud FROHMANN, 1983). Trata-se de um problema semântico porque a escolha das facetas requer uma decisão que exige que as práticas humanas sejam verificadas. Desde que não existam restrições *a priori* sobre o modo humano de empregar

¹¹ A garantia literária determina que o vocabulário escolhido para a construção de instrumentos terminológicos deve derivar empiricamente da literatura da área representada, ou seja, a literatura tem papel determinante para definir os conceitos que constituirão um determinado instrumento.

¹² No original, “thus, a basic problem in KO is whether semantic relations are *a priori* or *a posteriori*: whether they can be known before examining the literature or only after such an examination has been carried out. What kind of literary warrant (or other kind of warrant) is needed in order to identify semantic relations and classify concepts?” (HJORLAND, 2007c, p. 378)

palavras nas práticas linguísticas, não existiria outro modo para determinar relações semânticas entre termos que verificar como as pessoas empregam as palavras.

Pessoas comprometidas em um tipo de atividade podem empregar termos de uma forma, enquanto que aqueles comprometidos com atividades diferentes podem empregar os mesmos termos, de forma diferente. [...] Atividades humanas são fundamentais e determinam [o significado] preferencialmente que os significados pré-determinados das palavras¹³ (FROHMANN, 1983, p. 13).

Para a determinação das relações semânticas de um dado domínio, Hjørland (2007c) ressalta a necessidade de ‘garantias’ para validar o estabelecimento de relações. Além da garantia literária, outras garantias (garantia do usuário, garantia cultural, etc.) serão tratadas adiante, na seção 4.2, do presente trabalho. As relações semânticas são essenciais para a construção de ferramentas para a organização do conhecimento, os chamados sistemas de organização do conhecimento – SOC (tradução do inglês *knowledge organization systems* – KOS, de BRASCHER; CAFÉ, 2008). Neste trabalho, adotaremos a sigla SOC ao tratar estas ferramentas. Ponto comum a estes sistemas é o fato de se tratarem de ferramentas semânticas (*semantic tools*) e assim, comungarem das mesmas bases teóricas em semântica.

O termo *sistemas de organização de conhecimento* está pretendendo incluir todos os tipos de esquemas para organizar a informação e promover o gerenciamento do conhecimento. [...] Eles também incluem esquemas menos tradicionais, tais como redes semânticas e ontologias. Desde que os sistemas de organização de conhecimento sejam mecanismos para a organização de informação, eles [SOC] são centrais¹⁴ (HODGE, 2000, p. 3, tradução nossa).

¹³ No original, “[...] People engaged in one kind of activity may employ particular terms in one kind of concept organization, while those engaged in a different activity may employ the same terms, with the same references, in a different concept organization. Human activities are fundamental in this second view of language, and the activities determine rather than are determined by the meanings of words” [Na primeira abordagem, os relacionamentos são determinados *a priori* porque os relacionamentos se baseariam no significado dos termos envolvidos] (FROHMANN, 1983, p.13).

¹⁴ No original, “The term *knowledge organization systems* is intended to encompass all types of schemes for organizing information and promoting knowledge management. Knowledge organization systems include classification schemes that organize materials at a general level (such as books on a shelf), subject headings that provide more detailed access, and authority files that control variant versions of key information (such as geographic names and personal names). They also include less-traditional schemes, such as semantic networks and ontologies. Because knowledge organization systems are mechanisms for organizing information, they are at the heart of every library, museum, and archive” (HODGE, 2000, p. 3).

Segundo Hjørland (2007c, p. 368), os SOC apresentam conceitos selecionados mais ou menos enriquecidos com informações sobre suas relações semânticas e englobam desde listas de termos, como glossários e dicionários até as ontologias (ilustração 9).

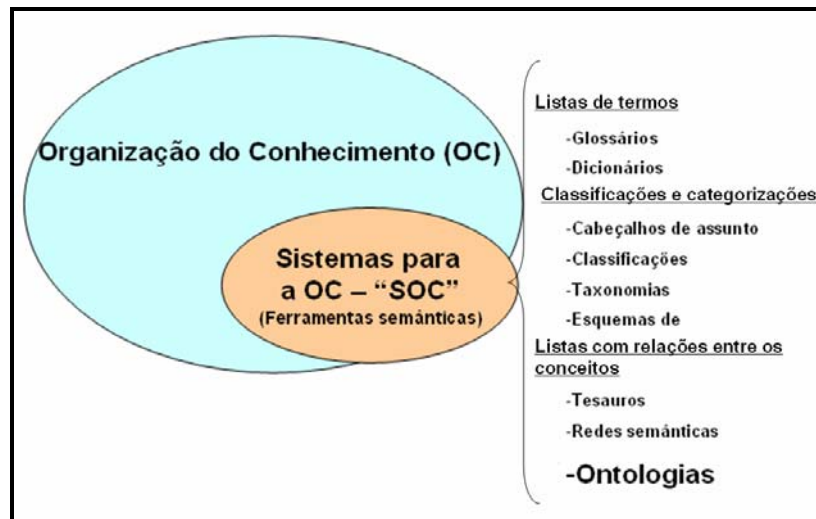


Figura 06 - Sistemas para a organização do conhecimento (Taxonomia de SOC)

Fonte: HODGE (2000). Adaptação da autora.

Para possibilitar os processos para a organização do conhecimento, segundo Souza (2000), ao longo do tempo os [SOC] foram se tornando cada vez mais sofisticados e complexos na tentativa de melhor representar o conteúdo intelectual dos documentos para fins de recuperação frente a necessidades cada vez mais complexas e diversificadas de recuperação de informação. As formas como se dão estes relacionamentos em SOC são enfatizadas nos próximos capítulos e orienta a presente pesquisa.

2.3 Relações semânticas em ciência da informação

Estudos sobre as relações apresentam-nas tanto como relações semânticas (FROHMANN, 1983; KHOO; NA, 2006; HJORLAND, 2007b; 2007c), como relações conceituais (DAHLBERG, 1995; MIRANDA, 2005; SALES, 2006; DODEBEI, 2002) ou ainda, como relações entre os

termos (MOTTA, 1987) sem maiores distinções sobre a nomenclatura. Neste trabalho, determinou-se tratá-las por relações semânticas, cujo propósito é aproximar nosso estudo aos projetos para o desenvolvimento da web semântica e do estudo do significado, propriamente.

Khoo e Na (2006) partem do princípio que conceitos e relações são fundamentos do conhecimento e do pensamento e, então, necessários para o raciocínio e inferências:

Relações semânticas são associações significativas entre dois ou mais conceitos, entidades ou conjunto de entidades. [...] Os conceitos/entidades são parte integral da relação, assim como uma relação não pode existir por ela mesma. Associações entre conceitos/entidades podem ser categorizadas em diferentes tipos, abstraídas, concebidas e diferenciadas de outras associações e podem portanto conferir significado (KHOO; NA, 2006, p. 159, tradução nossa)¹⁵.

Um ponto importante para o estudo das relações semânticas, resumido aqui, são as propriedades para as relações, citadas em algumas metodologias para o desenvolvimento de ontologias, principalmente ao que tange a formalização das linguagens. São elas: a simetria; transitividade e reflexividade, especificadas, a seguir:

- a) Simetria - quando a relação entre o conceito A e B é a mesma relação entre B e A. Em vocabulários controlados, relações são simétricas quando a relação usada é a mesma, independente da direção dos conceitos. A assimetria ocorre quando a relação entre o conceito A e B é inversa;
- b) Transitividade: uma relação R é uma relação transitiva se x relacionado a y e y relacionado a z , implica que x é relacionado a z . Se $a=b$ e $b=c$, então $a=c$ (SOWA, 2000, p. 95); e
- c) Reflexividade: uma relação R é uma relação reflexiva se a entidade se relaciona com ela própria. Para todos a , $a=a$ (SOWA, 2000, p. 95).

¹⁵No original, “Semantic relations are meaningful associations between two or more concepts, entities or sets of entities. [...] The concepts/entities are an integral part of the relation as a relation cannot exist by itself but have to relate two things. Associations between concepts/entities can be categorized into different types, abstracted, conceptualized and distinguished from other associations, and can thus be assigned meaning” (KHOO; NA, 2006, p. 159).

Mais que apresentar as diferentes terminologias para o nosso objeto, importa esclarecer o seu papel na representação e recuperação do conhecimento. Em SRIs tradicionais, existem instrumentos para auxiliar a entrada (representação) e a saída (recuperação) de documentos. O papel das relações semânticas, neste contexto, retoma os conceitos de revocação, precisão, exaustividade e especificidade – enquanto medidas para controle e avaliação do desempenho da recuperação em SRI, onde as relações semânticas têm papel crucial.

Dada a seguinte situação: dois conceitos, A e B, são os descritores selecionados através de uma ferramenta para representar o conteúdo de um determinado documento ou item informacional. Sem especificar a relação semântica entre os dois conceitos, há de se concordar que falsas associações entre os conceitos podem ser interpretadas (pelo usuário no momento da busca e pelo sistema ao processar os resultados da busca) e resultam em prejuízo para a revocação e precisão. A probabilidade de ocorrência de falsas associações desse tipo aumenta consideravelmente com a exaustividade da indexação. No entanto, mais que evitar falsas associações, a especificação das relações entre os conceitos representativos de um assunto está no fato que, se o assunto A é relacionado ao assunto B, então na busca, o usuário deve ser atentado sobre esta informação, independente do sistema. Maniez (1988) destaca que um usuário pode não estar totalmente certo de encontrar todos os documentos relevantes depois de uma primeira busca, quaisquer que sejam as qualidades da linguagem e das fórmulas utilizadas. As linguagens devem ter mecanismos de, a partir de uma fórmula de busca, apresentar outras associações possíveis para auxiliar na representação (no momento da indexação) e na recuperação (para ampliar a revocação, do inglês *recall*, para designar a capacidade de recuperar documentos úteis).

A precisão, segundo Maniez (1988), pode ser alcançada desde que os termos empregados pelo indexador e pelo usuário final em um dado sistema sejam precisos e inequívocos (a linguagem do sistema e a linguagem do usuário devem ser “compatíveis” ou não haverá comunicação), o que implica na rigorosa escolha dos descritores mais adequados para fins de representação - precisão designa a capacidade de evitar documentos inúteis.

Em Hjørland (2007c), para a recuperação da informação a função básica das relações semânticas é aumentar a revocação do sistema: “a inclusão de sinônimos e termos genéricos pode contribuir

para aumentar a revocação, enquanto a diferenciação de homônimos e a especificação de termos podem ampliar a precisão” (HJORLAND, 2007c, p. 392).

Já a especificidade, segundo a NBR 12.676, “se refere ao grau de precisão com que um termo define determinado conceito no documento. Ocorre perda de especificidade quando um conceito é representado por um termo com significado mais genérico” (ABNT, 1992, p. 3).

Em ambiente digital, a recuperação da informação passa a necessitar de maior atenção pela área, seja pela atualidade, seja pelo papel deste ambiente para a comunicação humana. Na web, não há o mesmo nível de controle passível dos SRI tradicionais, ao contrário, neste ambiente são aumentadas as possibilidades de associações devido a extensão do registro, como lembra Lancaster (2004, p. 189), devido a busca se dar principalmente em texto completo e a não-estruturação dos dados permitir inúmeras falsas associações entre os termos.

Em bases de dados muito grandes torna-se, progressivamente mais difícil alcançar um nível de revocação aceitável com um nível de precisão satisfatório, uma situação que chegou a um ponto crítico quando se procura informação na internet (LANCASTER, 2004, p. 4).

Mudado o ambiente, as ferramentas para o controle de vocabulário carecem de revisão e possíveis adaptações. Ontologias, por sua estrutura relacional mais aprimorada que ferramentas desenvolvidas anteriormente, ganham relevância para uso, desenvolvimento e pesquisa. Não se trata da evolução das primeiras, mas numa perspectiva histórica, ontologias se apresentam como ferramentas terminológicas com condições de agregar o maior número e variedade de relações entre os conceitos (sem dar margem à ambigüidade ou a falsas associações entre os conceitos) e por isso, também são ferramentas apropriadas para a organização do conhecimento.

Do ponto de vista da modelagem das ferramentas existentes, estas podem se dividir em dois tipos: ferramentas pré-coordenadas e pós-coordenadas. As relações semânticas são estabelecidas ou no desenvolvimento e construção da ferramenta (sistemas pré-coordenados) ou no seu uso (sistemas pós-coordenados), especificadas a seguir.

Em ferramentas pré-coordenadas, as relações são definidas *a priori* como parte do processo de sua construção. Jansen (1974 apud RIVIER, 1992) trata as ligações de um termo com todos os

outros descritores pré-coordenados por relação de dependência – que não deve ser confundido com relações hierárquicas. O método de indexação altamente estruturado (com o emprego de elos, funções ou códigos semânticos) predominou no início da década de 1960, quando os sistemas informatizados ainda se achavam num estágio de desenvolvimento muito preliminar (LANCASTER, 2004, p. 195).

Nos sistemas pré-coordenados, segundo Lancaster (2004, p. 50), é difícil representar a multidimensionalidade das relações entre os termos - que podem ser listados numa determinada seqüência (A, B, C, D, E), onde o primeiro termo é mais importante do que os outros e, não é fácil (senão completamente impossível) combinar termos no momento em que se faz uma busca.

Os sistemas pós-coordenados, por sua vez, surgiram na década de 1940, quando foram implantados com a utilização de vários tipos de fichas. Um sistema informatizado moderno, funcionando em linha, pode ser visto como um descendente direto desses sistemas manuais (LANCASTER, 2004, p. 38). Neste sistema, os termos podem ser combinados entre si de qualquer forma no momento em que se faz a busca; preserva-se a multidimensionalidade das relações entre os termos e todo termo atribuído a um documento tem peso igual - nenhum é mais importante do que outro.

Apreendidas as funções, propriedades e coordenação das relações semânticas, pode-se afirmar que ontologias são ferramentas que podem incorporar, no seu desenvolvimento, a estruturação explícita e *a priori* das relações, como em sistemas pré-coordenados, sem abster da flexibilidade de navegação possível em sistemas pós-coordenados. A partir desta afirmativa, segue a apresentação das relações semânticas presentes nas ferramentas para a OC em ciência da informação, no capítulo a seguir.

2.3.1 Relações semânticas em instrumentos terminológicos

A pesquisa sobre relações semânticas em ciência da informação tem longo histórico, como verificado na literatura da área. Para isso, buscou-se destacar os tipos de relações apresentadas na

literatura. Este levantamento partiu da pesquisa sobre os tipos de relações em ferramentas para a organização do conhecimento e chegou aos seguintes instrumentos e projetos, ordenados cronologicamente: sistema unitermo (1953); projeto Syntol (1968); sistema PRECIS (AUSTIN, 1974), sistemas de classificação bibliográfica (CDD, CDU, *Colon Classification*), indexação relacional (FARRADANE, 1980b; 1980c); tesouros (MOTTA, 1987; GOMES, 1990), teoria do conceito (DAHLBERG, 1978b), explorados no presente capítulo.

O sistema unitermo é um sistema pós-coordenado, que consiste na indexação por palavras únicas retiradas do contexto, na maioria das vezes, sem controle de vocabulário, criado e empregado por Mortimer Taube em 1953 (VALE, 1987, p. 20). Neste sistema, as palavras eram utilizadas como unidades de representação do conteúdo dos documentos, mas podiam ser combinadas através da comparação das fichas referentes aos termos que se está procurando para a recuperação mais a indicação do número referente ao documento: todo documento recebe um número de registro, em ordem crescente e independente do assunto tratado, chamado número de acesso (FOSKETT, 1973, p. 311).

Com a introdução do computador, no Projeto Syntol (Syntagmatic Oriented Language), Cros; Gardin e Levy (1968) visavam uma indexação sintática auxiliada por computador. Este projeto trabalhava com a distinção explícita entre as relações *a priori* (relações paradigmáticas) e as relações *a posteriori* (relações sintagmáticas). Gardin, no entanto, afirma que esta distinção depende da forma como são estabelecidas dentro do sistema, mas não em essência (GARDIN, 1973, p. 14).

Maniez (1988, p. 133) compreende que a distinção sintagmática-paradigmática é essencial por diferenciar duas operações mentais: comparar e relacionar. Segundo Cunha (1987, p. 45),

a língua funciona segundo dois eixos (estudados pela gramática): o sintagmático, que rege os agrupamentos possíveis dos diferentes signos lingüísticos e o eixo paradigmático, que reflete as relações existentes entre os signos capazes de assumir a mesma função. [...] No caso da análise documentária, que tem como objetos textos, isto é, enunciados finitos e espaços discursos limitados, o reconhecimento das relações sintagmáticas e paradigmáticas leva não só à formulação de campos semânticos dos textos analisados como à estruturação e elaboração das linguagens tradutoras, permitindo a passagem de uma LN [linguagem natural]- LD [linguagem documentária] (CUNHA, 1987, p. 45).

No Sistema PRECIS (*Preserved Context Indexing System*), de Austin (1974, segundo Foskett, 1973, p. 62) faz-se a organização sintática de enunciados de assunto e utiliza-se operadores relacionais inseridos entre os conceitos. O avanço que o PRECIS representa é permitir maior precisão quanto ao tipo de relação entre os conceitos e, por conseguinte, possibilitar a elaboração de enunciados básicos, também processáveis por computador. No entanto, a importância de citar o Sistema PRECIS (Quadro 03) está em trabalho posterior: a indexação relacional de Farradane – tratada adiante.

Quadro 03 - Operadores do PRECIS

- (a) Forma; física, p. ex. microfilme, ou bibliográfica, p. ex. revista
- (b) Alvo; p. ex. para bibliotecários
/ Filiação de campo, relação quase-genérica
, Diferença
- (p) Subsistema, estrutura, material
- (q) Propriedade
- (o) Região de estudo, amostragem da população
- (1) Ponto de vista, perspectiva
- (2) Sistema ativo
- (3) Efeito, ação
- (4) Sistema-chave
- (5) Disciplina
- (6) Ambiente
- (v) Conceito coordenado
- (w) Assunto coordenado correlato
- (x) Assunto coordenado no mesmo documento

Fonte: FOSKETT, 1973, p. 63; adaptado pela autora.

Os sistemas de classificação bibliográfica objetivam a guarda física de documentos em bibliotecas e unidades de informação, de forma que os itens co-relacionados fiquem fisicamente próximos. Este tipo de arranjo permite o *browsing*, onde o usuário ao identificar a notação do assunto de interesse, pode ampliar a recuperação de documentos úteis relacionados (taxa de revocação). As relações em sistemas de classificação bibliográficas são hierárquicas, tendo como única relação a relação de classe-inclusão ou multidimensional (neste caso, engloba as relações parte-todo; instrumento-alvo, etc. sem explicitar suas particularidades).

Dos principais sistemas, três podem ser citados:

- a) CDD (Classificação Decimal de Dewey, criada em 1876): divide o conhecimento humano em 10 classes.
- b) CDU (Classificação Decimal Universal): Seus mentores foram *Paul Otlet e Henri La Fontaine*, em 1905. Este esquema também divide o conhecimento em 10 classes e é possível expressar (por símbolos) não apenas os assuntos simples, como também as relações entre os diversos assuntos (SOUZA, 2004, p. 27). Por exemplo, procura-se especificar qualquer combinação de conceitos através do emprego das tabelas auxiliares à tabela principal para atingir um grau maior de especificidade e de recuperação de assuntos para possibilitar a ordenação de documentos em estantes, principalmente. Para isso, este sistema faz uso de sinais de adição (+) para representar a assuntos coordenados; a barra oblíqua (/) para classes seqüenciais (dentro das classes propostas no sistema, representa extensão); dois pontos (:) para indicar que as classes se relacionam; colchetes ([]) para indicar subagrupamento e os dois pontos duplos (::), processável por máquina, que indica uma não-relação; entre outras, apenas para enumerar algumas; e
- c) Classificação de dois pontos (*Colon Classification* de Ranganathan, criada em 1933). Neste sistema é possível relacionar as características dos diversos assuntos. Segundo Souza (2004, p. 18), sua principal característica é a subdivisão dos assuntos em facetas e focos. Ex.: na subdivisão do assunto VESTUÁRIO, sua faceta COR possui vários focos: VERDE, AZUL, AMARELO, etc.

Ainda sobre a Classificação de dois pontos, Ranganathan (1967 apud CAMPOS, 2001), insatisfeito com as classificações da época, elabora seu próprio esquema de classificação e desenvolve a teoria de análise de facetas, demonstrando que a análise e a síntese se aplicam a qualquer classe básica e podem ser sistematizadas (FOSKETT, 1973, p. 264). Os problemas em relação à síntese em uma linguagem notacional levaram-no a desenvolver uma de suas teorias mais importantes, o PMEST: Personalidade, Matéria, Energia, Espaço e Tempo, cuja importância extrapola as cinco facetas propostas, mas as facetas em si.

Há cinco e somente cinco Categorias Fundamentais, são elas: Tempo, Espaço, Energia, Matéria e Personalidade. Estes termos e as idéias denotadas são usadas estritamente no

contexto da disciplina de classificação. Não têm nada a ver com seu emprego em Metafísica ou Física. Em nosso contexto, seu significado pode ser visto somente nas declarações sobre as facetas de um assunto – sua separação e seqüência. Este conjunto de categorias fundamentais é, em síntese, denotado pelas iniciais PMEST (RANGANATHAN, 1967, p. 398 apud CAMPOS, 2001, p. 56).

Para Farradane (1955) o sistema mais desenvolvido é o *Colon Classification*, onde os termos obtidos por análise de um assunto complexo são alocadas em uma das cinco categorias de conceitos. No entanto, a natureza das relações é indefinida. Para ele, a descoberta e o reconhecimento das relações entre conceitos são fundamentais para o aprendizado e estas não podem ser negligenciadas. Farradane (1952, 1955, 1980b; 1980c) desenvolveu a indexação relacional, que adota operadores relacionais, para preservar a semântica entre os conceitos indicada pelo indexador no momento da indexação, também no momento da recuperação – em um método apropriado para o processamento por máquinas. Defendia a necessidade de pré-coordenação dos termos e apresentava críticas aos sistemas de classificações bibliográficas - por serem muito estáticas e aos tesouros por apresentarem os termos desconectados, limitados às relações que pretendem orientar o indexador e o usuário.

De acordo com Farradane, nove relações podem ser encontradas na prática e são suficientes para os propósitos da indexação. Antes de apresentar a indexação relacional, faz-se necessário esclarecer o que é o conceito de indexação, em ciência da informação. Indexação é o processo de atribuir conceitos que sintetizem e representem o conteúdo de um documento em um sistema de recuperação da informação. O UNISIST (1981) considera a indexação enquanto processo que consiste em descrever e identificar um documento com a ajuda de representações dos conceitos nele contidos e quanto à sua finalidade, permitir busca e acesso a informação armazenada.

Farradane se preocupou em explicitar as relações entre os conceitos para que o significado pretendido no momento da indexação não se perdesse no momento da recuperação da informação. Segundo o autor, os conceitos podem variar muito caso a caso, mas afirma que esta

estrutura de relações invariável e pode expressar relações de aplicação geral em qualquer área de assunto e em qualquer nível de complexidade¹⁶ (FARRADANE, 1980b, p. 267).

Na indexação relacional, Farradane define relações, ou mais corretamente, categorias de relações para associar pares de conceitos e expressar as relações com base nos mecanismos do pensamento, convertidos diretamente em uma notação de indexação, processável pelo computador. Baseado na psicologia infantil, seu sistema compreende que o processo de aprendizagem começa pelo desenvolvimento de faculdades de discriminação no tempo e no espaço. No tempo, a primeira etapa é o ‘não-tempo’ (a ocorrência simultânea de duas idéias sem referência ao tempo); a segunda é a não-tempo (a ocorrência simultânea de tempos em tempos, mas não permanentemente, de duas idéias) e a terceira é fixa (a ocorrência simultânea e permanente entre duas idéias). No espaço, as etapas de discriminação são: a primeira é a coincidente (dois conceitos difíceis de discriminar um do outro); a segunda é a não-distinta (dois conceitos que tem muito em comum) e a terceira é a distinta (dois conceitos que se podem diferenciar completamente). Estes dois conjuntos de graduações em Farradane formam uma matriz em que os pontos de interseção denotam nove tipos de relações.

A teoria de Farradane está sintetizada no Quadro 04, a seguir. No cabeçalho horizontal do quadro, temos os três estágios de associação: a consciência (relação não-tempo), a associação temporária e a associação fixa. Verticalmente, com os três estágios de discriminação: a coincidente, a não-distinta e a distinta.

¹⁶ No original, “Relational indexing, as described here, is a means of expressing relations on the basis of the mechanisms of thought, to be converted directly into indexing notation. The subject concepts involved may be very different from case to case, but the framework of possible relations is limited and invariant, and thus bypasses the subject field and linguistic problems, and can be applied at any level of detail or complexity” (FARRADANE, 1980b, p. 267).

Quadro 04 - Categorias relacionais de Farradane

CATEGORIAS RELACIONAIS	Associação		
	Consciência	Associação temporária	Associação fixada
Discriminação			
Conceitualização concomitante	Concomitância / θ	Auto atividade /*	Associações /;
Conceitualização não distinta	Equivalência /=	Dimensional /+	Pertencimento /()
Conceitualização distinta	Distinção /)	Ação /-	Causação /:

Fonte: FARRADANE, 1980b, p. 270.

As nove relações, propostas por Farradane (1980b) são enumeradas a seguir:

- Relação de concomitância ou coincidência (*concurrency*): Seriam dois conceitos difíceis de discriminar, cuja percepção é simples, sem referência ao tempo. Implica uma mera justaposição mental de dois conceitos. Coincidência ou justaposição (mental) de uma idéia/objeto com outro, por exemplo: A na presença de B. Usado quando dois conceitos são indistinguíveis, independente do tempo. Ex.: Britânica / θ Enciclopédia (FARRADANE, 1980b, p. 270).
- Relação de equivalência (*equivalence*): quando dois conceitos têm muito em comum, cuja percepção é simples, sem referência ao tempo. Dá-se pela expressão de algum grau de equivalência ou equivalência completa entre conceitos (como o caso do sinônimo). Trata-se de uma relação permanente e indistinta. Ex.: sódio/=íon; melaços de cana/= forragem; acetona /= solvente (FARRADANE, 1980b, p. 270).
- Relação de distinção (*distinctness*): quando dois conceitos podem ser completamente discriminados, cuja percepção é simples, sem referência ao tempo. Aplica-se para expressar relação de imitação ou substituição. É atemporal, mas ao contrário da relação de co-ocorrência, é uma relação que expressa a distinção entre os conceitos. Ex.: homem /) estátua; recuperação da informação /) modelo matemático (FARRADANE, 1980b, p. 270).

- Relação de comparação ou auto-atividade (*self-activity*): Quando dois conceitos são difíceis de discriminar e cuja associação entre os dois é temporária. Associa-se a verbos intransitivos como migrar e caminhar. Segundo o autor, é de difícil avaliação (*assess*) teórica, mas muito útil na prática. Ex.: pássaro /* migrando; homem /* caminhando (FARRADANE, 1980b, p. 271).
- Relação dimensional e estado (*dimensional*): quando os conceitos têm muito em comum e a associação entre eles é temporária. Expressa posição no espaço ou tempo, estados temporários e propriedades temporárias. Relaciona os estados ou propriedades temporárias de um conceito ou determinada classe (estes estados temporários incluem temperatura, forma, bem como propriedades variáveis de tamanho, peso, volume, entre outros). Ex.: sal /+ cristal; livro /+ estante (FARRADANE, 1980b, p. 271).
- Relação ação (*action*): quando os dois conceitos podem ser completamente discriminados e a associação entre os dois é temporária. Esta relação pode ser usada para descrever uma operação ou o efeito (“afetação”) de uma coisa sobre outra. Ex.: água /- purificação; decomposição /- prevenção; roupas /- detergente (FARRADANE, 1980b, p. 271).
- Relação associação (*association*): Quando dois conceitos são difíceis de discriminar, cuja associação entre eles é fixa, permanente. Esta relação expressa várias formas de associação, as quais podem ser indefinidas (prisão /; desgraça), ou a relação com um agente de, ou ferramenta para, uma operação, etc. Expressa propriedades abstratas (obra de arte/; beleza) ou relações de ações no passado ou não especificadas (FARRADANE, 1980b, p. 271).
- Relação de pertença (*appurtenance*): quando os conceitos têm muito em comum e cuja associação entre eles é fixa, permanente. Expressa relação de todo-parte, de gênero-espécie e toda propriedade física intrínseca de um material ou coisa particular. Ex.: mesa /(perna; gênero /(espécie; chá /(cafeína. Expressa ainda, relação para todas as propriedades físicas e intrínsecas de um objeto em particular (por exemplo: metal /(

densidade; tubo / (diâmetro) (FARRADANE, 1980b, p. 271).

- Relação dependência funcional (*functional dependence*): quando dois conceitos podem ser completamente discriminados, cuja associação entre eles é fixa, permanente. Expressa a relação de uma coisa causando ou produzindo algo ou mesmo um produto produzido por outra ‘coisa’. Aplica-se, especialmente, para a indexação de reações químicas. Ex.: trigo /: pão; autor /: livro (FARRADANE, 1980b, p. 272).

O sistema proposto por Farradane sofreu críticas devido à sua generalidade (por exemplo, as relações de Equivalência e de Associação) e quanto à dificuldade de consenso entre indexadores (devido à subjetividade das relações) dentro de um mesmo domínio, o que demonstra que o estabelecimento de relações pode não ser tão consensual/cognitivo (MYAENG; McHALE, 1991; FOSKETT, 1973).

Myaeng e McHale (1992) apontaram dois problemas com as relações de Farradane: a discrepância na granularidade e a ambigüidade das relações. A relação de equivalência implica que um sinônimo ou uma expressão equivalente é útil para alguns poucos itens. A relação de associação, por outro lado, pode ser usado para expressar qualquer relação específica. E este é um problema porque a relação de associação pode ser usada mais que qualquer outra relação e ainda que se remova a relação de equivalência, ainda assim, podemos classificar todas as relações porque podemos usar a associação para uma equivalência não-especificada.

Um exemplo da aplicação da teoria é apresentado em Brookes (1986), onde podemos identificar os pares de conceitos e respectivas relações.

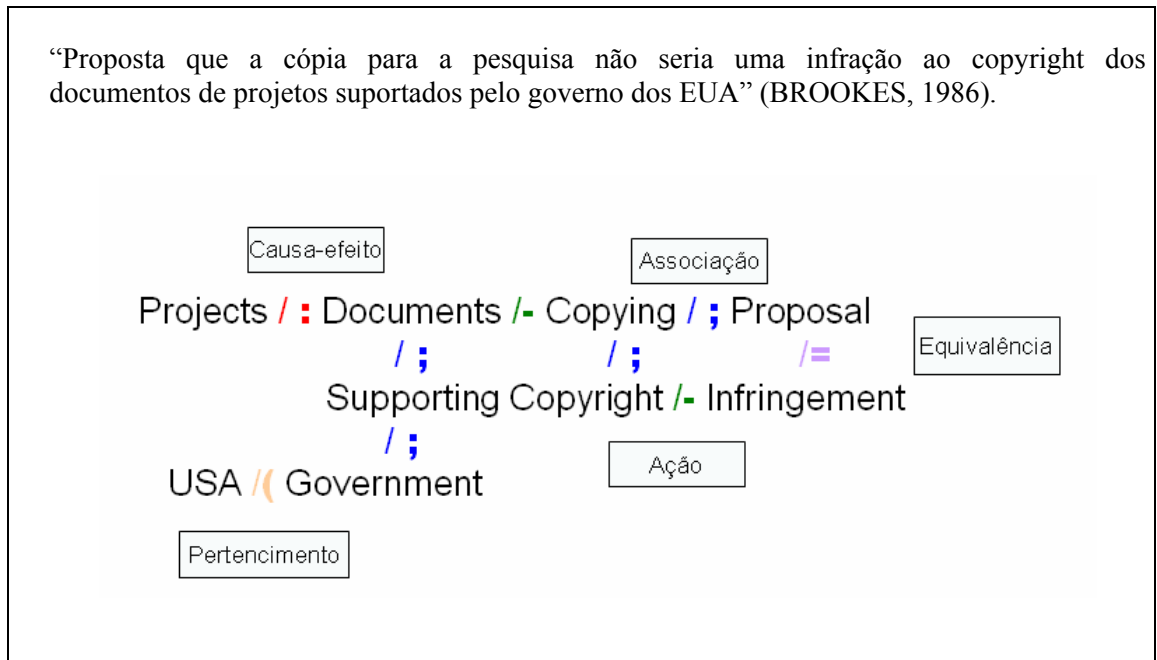


Figura 07 - Exemplo da aplicação da indexação relacional de Farradane.

Fonte: BROOKES (1986).

Consideremos que no processo de distinção das relações existentes, Farradane utiliza simultaneamente dois processos contrários – associação e discriminação, similares ao próprio processo de classificação, que segundo PIEDADE (1983) é definido como “dividir em grupos ou classes, segundo as diferenças e semelhanças. É dispor os conceitos, segundo suas semelhanças e diferenças, em certo número de grupos metodicamente distribuídos” (PIEADADE, 1983).

Em outro momento da literatura, observamos um avanço na análise da natureza das relações: a teoria do conceito, desenvolvida pela filósofa Ingetraut Dahlberg (1978a; 1978b). Para Dahlberg (1978b, p. 104) “[...] sempre que diferentes conceitos possuem características idênticas, deve-se admitir que entre eles existam relações”.

Um conceito é uma unidade de conhecimento, abrangendo afirmações sobre um item selecionado ou referência, representado em uma forma verbal. [...] uma afirmação verificável é o componente de um conceito o qual situa um atributo do item dele de referência. [...] um item de referência é o componente do conceito para o qual suas afirmações verificáveis e sua forma verbal estão diretamente relacionadas, assim seu “referente”. [...] uma forma verbal (termo/nome) de um conceito é o componente no qual

convenientemente sumariza ou sintetiza e representa um conceito do propósito de comunicação¹⁷ (DAHLBERG, 1978a, p. 143-144).

Dahlberg (1979) faz um levantamento das pesquisas sobre as relações entre conceitos, e constata que com o aumento da literatura documentada e a insatisfação com a qualidade da recuperação da informação, ressurgiu o interesse sobre este campo na ciência da informação. Afirma ainda que, os tesouros muito se desenvolveram graças à necessidade real de representação de assuntos mais específicos que as notações dos sistemas de classificações tradicionais.

Através do exercício intelectual de tentar identificar as relações entre termos que foi criada uma nova compreensão da estrutura conceitual do conhecimento, uma vez que um tesouro é definido como uma lista de conceitos e termos correspondentes numa determinada área do conhecimento, com a indicação das relações entre os conceitos e os termos que denotam (DAHLBERG, 1979).

Dahlberg (1978b, p.104; 1978a, p. 148) distingue dois tipos principais de relacionamentos: quantitativo e qualitativo. A relação quantitativa mede o nível de similaridade das características de um conceito:

- a) Identidade: As características do conceito A (x,x,x) e do conceito B (x,x,x) são as mesmas;
- b) Inclusão: As características do conceito A (x,x) estão contidas no conceito B (x,x,x);
- c) Intersecção: As características do conceito A (x,x,o) coincidem em algum elemento com as características do conceito B (x,o,o); e
- d) Disjunção: Nenhuma característica em comum em A (x,x,x) e B (o,o,o).

Qualitativamente, Dahlberg (1978b, p. 104-105) divide as relações em: hierárquicas, partitivas, de oposição e funcionais.

¹⁷No original, “A concept is a knowledge unit, comprising verifiable statements about a selected item of reference, represented in a verbal form. [...] a verifiable statement is the component of a concept which states an attribute of its item of reference. [...] a item of reference is the component of the concept to which its verifiable statements and its verbal form are directly related, thus its “referent”. [...] a verbal form (term/name) of a concept is the component which conveniently summarizes or synthesizes and represents a concept for the purpose of designating a concept in communication” (DAHLBERG, 1978a, p. 143-144).

- a) Relações hierárquicas (implicação): a relação entre gênero e espécie; espécie e espécie; espécie e indivíduo.
- b) Relações partitivas: Relação entre um todo e suas partes; entre as partes e subpartes.
- c) Relação de oposição (negação): podem ser de contradição ou de contrariedade.
- d) Relações funcionais (intersecção): aplicam-se sobretudo a conceitos que expressam processos. Pode-se conhecer o caráter semântico de tais relações tendo por base as chamadas valências semânticas dos verbos, dando atenção aos verbos e respectivos complementos. Ex.: ‘produzir’ demanda um produtor, um consumidor, um produto, etc.

A Teoria Geral da Terminologia - TGT de Wuester (1981 apud RIVIER, 1992) tem por objetivo racionalizar o processo da comunicação

[...] seu trabalho assemelha-se ao de elaboração de um tesouro: deve-se identificar em uma área do saber, primeiramente os conceitos, depois lhe atribuir um termo (em uma ou em diversas línguas) controlando, ao mesmo tempo, suas relações de sinonímia, homonímia, etc. seguindo ainda, um método dedutivo para dar uma definição rigorosa de cada conceito (RIVIER, 1992, p. 80).

Segundo Campos (2001, p. 74) as relações tratadas em nível conceitual se dividem em relações lógicas - resultam forçosamente da própria compreensão dos conceitos (relação de abstração, de semelhança) e ontológicas - relações indiretas entre conceitos, porque resultam das propriedades que possuem os representantes dos conceitos (os objetos do mundo empírico - o conceito e a realidade). Sendo que as relações lógicas são: relação de determinação, de conjunção e disjunção; e as relações ontológicas são as relações de contato (relações de coordenação e de encadeamento) e de causalidade (geral, filogênico, ontogênico, substâncias).

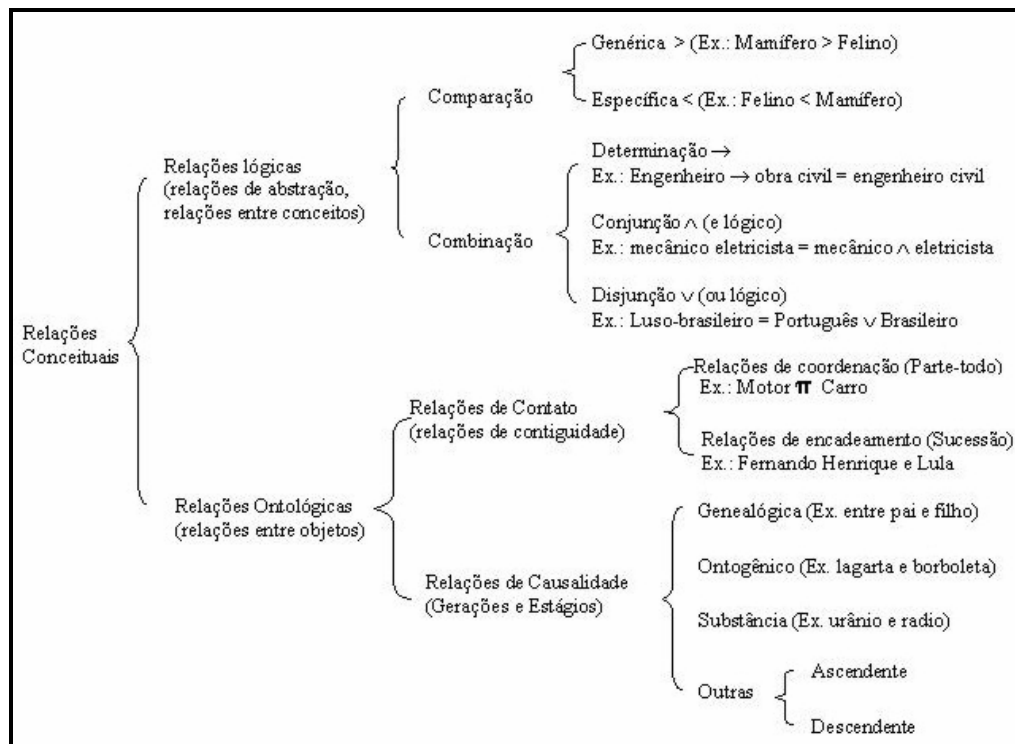


Figura 08 - Classificação das relações segundo a TGT.

Fonte: Moreira, 2005.

Campos (2000, p. 124) afirma que as relações lógicas e ontológicas não possuem a finalidade de estabelecer certa ordem entre os conceitos, mas sim de determinar a natureza das relações que ocorrem entre eles. Para a presente pesquisa, optou-se por utilizar a distinção lógica *versus* ontológica para a determinação da natureza das demais relações, apenas. Moreira (2005) traça a seguinte relação entre as propostas de Wuester (1981) e a *International Standard Organization – ISO 704* (2000):

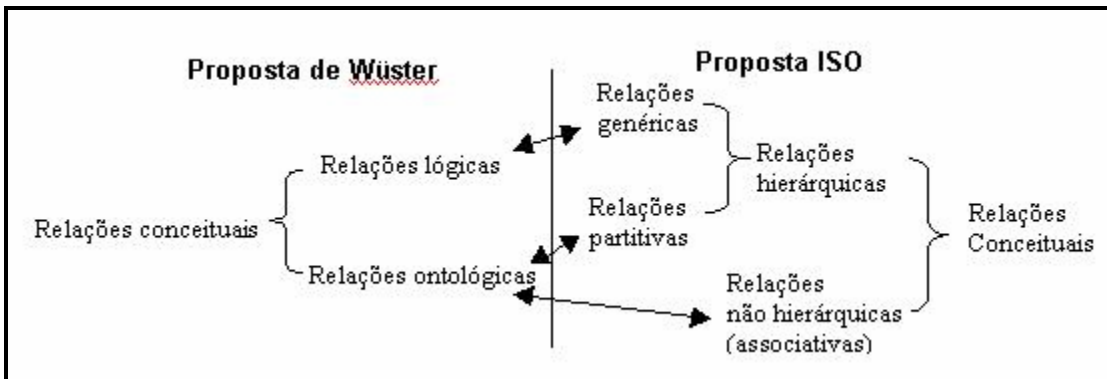


Figura 09 - Equivalência entre as relações conceituais.

Fonte: Moreira (2005)

A norma ISO 704 (2000, p. 5) divide as relações em: relação hierárquica (relação genérica e relação partitiva) e relação associativa.

Uma relação genérica ocorre entre dois conceitos quando a intensão¹⁸ do conceito subordinado (conceito específico) inclui a intensão do conceito superordenado (conceito genérico) mais uma característica adicional (ISO 704, 2000, p. 5).

Uma relação partitiva ocorre quando o conceito superordenado representa um inteiro, enquanto o conceito subordinado representa parte deste inteiro. As partes, juntas, formam um inteiro. O conceito superordenado em uma relação partitiva é chamado o conceito compreensivo (*comprehensive concept*) e o conceito subordinado é chamado conceito partitivo. Conceitos subordinados em um mesmo nível, que compartilham a mesma dimensão são chamados conceitos coordenados (ISO 704, 2000, p. 9).

A relação partitiva e a relação genérica podem ser expressas em séries verticais e horizontais. As partes de um inteiro podem ser de natureza similar (átomo e molécula de oxigênio) ou diferentes entre si. Uma ou mais partes podem ser essenciais ou não-essenciais (opcionais). Para identificar as características essenciais dos conceitos partitivos é necessário determinar a intensão do conceito compreensivo (*comprehensive concept*) anteriormente. Segundo a norma, um sistema conceitual partitivo nem sempre permite uma análise completa de um conceito. Se o conceito partitivo não é particular ao conceito compreensivo, então a extensão de um conceito partitivo

¹⁸ Trataremos a definição de intensão (DAHLBERG, 1978a) na metodologia do presente trabalho.

não explica completamente e as características de sua intensão podem se perder. Um conceito partitivo pode ser definido com base de uma relação partitiva apenas se a extensão completa e as características da intensão puderem ser determinadas (ISO 704, 2000, p. 10).

Já as relações associativas são relações não hierárquicas. Existe uma relação associativa quando uma conexão temática pode ser estabelecida entre conceitos em virtude da experiência. Algumas relações associativas ocorrem quando a dependência é estabelecida entre conceitos com respeito a sua proximidade no espaço ou no tempo. Algumas relações envolvem eventos no tempo, como a dependência de um processo no tempo ou seqüência; outras relações envolvem causa e efeito (ISO 704, 2000, p. 10). Estas relações podem envolver:

- a) Matéria prima/produto;
- b) Ação-equipamento/ferramenta;
- c) Quantidade/unidade;
- d) Material/propriedade;
- e) Material/estado;
- f) matéria/substância-propriedade;
- g) item concreto/material;
- h) item concreto/forma (*shape*);
- i) ação/alvo;
- j) ação - lugar/localização;
- k) ação/ator;
- l) recipiente/o que é contido (pencil case/pencil); (estojo, lápis);
- m) atividade/ferramenta (*tool*) (escrever; lápis);
- n) passos de um ciclo (gametas; zigoto; zigosporo);
- o) causa/efeito (umidade/corrosão);
- p) produtor/produto (padeiro; pão);
- q) duração/equipamento de medida (tempo; relógio);
- r) profissão/ferramenta típica (pintor; pincel);
- s) objeto/ferramenta associada (parafuso; chave de fenda);
- t) organização/edificação associada (*Muslim; mosque* muçulmano; mesquita) , etc.

Tesauros, segundo a UNESCO (1973 apud CAMPOS, 2001, p. 91) têm por função ser um instrumento de controle terminológico usado para traduzir a linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema mais restrita. Segundo a estrutura, é constituído por um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.

Para Clarke (2001):

tesauros em ambiente de controle de vocabulário é a ferramenta que suporta a efetiva recuperação da informação, guiando indexadores e pesquisadores (searchers) consistentemente para a escolha dos mesmos termos para expressar um dado conceito ou combinação de conceitos. Os termos em tesauros são relacionados por três tipos de relações bem-conhecidos: equivalência; hierárquica e associativa (CLARKE, 2001, p. 37)¹⁹.

Segundo Campos (2000, p. 91), “a evolução histórica do tesouro de recuperação [...] pode ser traçada a partir de duas vertentes: uma que toma nitidamente como base o unitermo [...] e a outra influenciada pela Teoria da Classificação Facetada.” Os primeiros, elaborados na América do Norte, a a partir da Segunda Guerra Mundial, com o desenvolvimento científico e tecnológico, a literatura passou a exigir um tipo de representação que os sistemas unitermos, nem os cabeçalhos de assunto conseguiram acompanhar.

A norma proposta pela organização norte-americana *National Information Standards Organization* (ANSI, 2005) propõe as linhas gerais para a construção, formatação e manutenção de vocabulários controlados monolíngües. De origem americana, é baseado no termo e não no conceito, por isso não há indicação de metodologia para a definição conceitual dos termos. Este documento define vocabulário controlado segundo uma ordem conhecida e estruturada com o intuito de disponibilizar claramente os relacionamentos de equivalência, associativos e hierárquicos, determinando desde níveis de relacionamentos simples até estruturas mais complexas. Divide, então, as relações em três: relação de equivalência, hierárquica e associativa.

¹⁹ No original, “a thesaurus in the controlled vocabulary environment is a tool designed to support effective information retrieval (IR) by guiding indexers and searchers consistently to choose the same terms for expressing a given concept or combination of concepts. Terms in the thesaurus are linked by relationships of three well-known types: equivalence, hierarchical, and associative” (CLARKE, 2001, p. 37).

A relação de equivalência é aquela que representa os sinônimos ou quase sinônimos de um termo. Divide-se em: relações de sinonímia; as variações lexicais entre conceitos e a quase-sinonímia.

A relação hierárquica engloba a relação do tipo genérica (gênero-espécie); a “Instance Relationships”, relação de instanciação (é um) e a relação do tipo partitiva (todo-parte).

A relação associativa, segundo a norma ANSI (2005, p. 51) é a mais difícil de definir e salienta que é importante explicitar a natureza das relações entre os termos e evitar julgamentos subjetivos, sob a pena de haver inconsistências. Podem ocorrer entre termos de uma mesma hierarquia ou entre hierarquias diferentes:

- a) causa/efeito;
- b) processo/agente;
- c) processo/contra-agente;
- d) ação/produto;
- e) ação/propriedade;
- f) ação/alvo;
- g) conceito ou objeto/propriedade;
- h) conceito ou objeto/origem;
- i) conceito ou objeto/ unidade ou mecanismo de medida;
- j) matéria bruta/produto e
- k) disciplina ou campo/objeto ou profissional (ANSI, 2005, p. 42).

Já na vertente europeia – segundo Rivier (1992, p. 72) e Campos (2000, p. 98), até o final dos anos 50 e início dos anos 60, os tesouros possuíam somente arranjo alfabético, tipo de arranjo que era incapaz de representar bem as relações úteis entre os termos. Limitações estas que levaram ao emprego de princípios de classificação.

O tesouro mostra as relações hierárquicas (TG/TE, termo genérico e termo específico) e os termos relacionados de outro modo que não o hierárquico (TR, termo relacionado), por exemplo, para os relacionamentos todo/parte ou processo/equipamento, e coisa/propriedade [...] (CAMPOS, 2000, p. 98).

Motta (1987) considera a problemática envolvida no embasamento teórico para a determinação das relações em tesouros. Nas palavras da autora:

mesmo se tratando de relações de gênero/espécie ou todo/parte, que podem ser mais facilmente identificadas, geralmente o estabelecimento dessas relações assume cunho altamente pessoal, opinativo, quer pela falta de definições conceituais corretas, quer pela falta de conhecimento que os conceptores (*sic*) de tesouros têm a respeito da área de assunto coberta pelo sistema (MOTTA, 1987, p. 39).

Em seu trabalho, Motta revê a literatura pertinente às relações semânticas, particularmente a relação associativa, já que o objetivo de seu trabalho é a construção de vocabulários. O resultado foi apresentado pela autora associando a cada par de conceitos (de um tesouro de Economia), o esclarecimento do tipo de relação associativa. Chega às possíveis formas para as relações associativas:

- a) Relação de causalidade: segundo Motta (1987, p. 48), de acordo com o exemplo: Crescimento Econômico TR Desenvolvimento Econômico, se inferiu a relação de causalidade por se considerar que “Desenvolvimento Econômico” não ocorre, sem que haja “Crescimento Econômico”; que, embora “Crescimento Econômico” não seja condição suficiente para o “Desenvolvimento Econômico”, ele é condição necessária.
- b) Relação instrumental: “qualquer coisa que usa, ou que é usada, tendo em vista um determinado fim, independentemente de produzir uma ação efetiva”(MOTTA, 1987, p. 48). Ex.: Política monetária TR Taxas de Juros (instrumento de política monetária).
- c) Relação de afetação ou de influência: “aquelas [relações] onde os conceitos exercem influencia um sobre o outro, incondicional ou mútua, sem que necessariamente se atribua as categorias de causa e efeito a um e a outro” (MOTTA, 1987, p. 48). Ex.: Política monetária [que influencia] Inflação; Planejamento econômico[que influencia] Alocação de recursos
- d) de oposição: Ex.: Emprego TR Desemprego.
- e) Relação interfaceta: “esta relação [interfaceta] se dá quando o termo B é associado a A porque A já havia sido associado a B previamente, pelo fato de A

ser uma das características de B, e sem que seja, necessariamente, uma das características de A” (MOTTA, 1987, p. 49). Ex.: Nível de atividade econômica TR Política monetária – especificamente neste exemplo, pertencem a duas facetas distintas, ou seja, Política monetária (da faceta instrumentos) e Nível de atividade econômica (faceta coisa)

- f) atributiva/atribuição: Relação identificada através de atributos que as coisas possuem (MOTTA, 1987, p. 50). Ex.: Economia TR Nível de atividade econômica.
- g) associação implícita: “relação que não se conformou com qualquer um dos anteriormente referidos, foi também identificado. Na falta de um nome mais adequado, tal relação foi chamada de ‘associação implícita’”(MOTTA, 1987, p. 50).

Nota-se que a relação de associação permanece. Segundo Dodebei (2002, p. 96), tal identificação é resultado pragmático da análise conceitual dos termos levantados, específicos para o campo da Economia e pode não se aplicar a outras áreas.

Quadro 05 – Extrato do tesouro de economia

DESEMPREGO	
TA DESENVOLVIMENTO ECONÔMICO	interfaceta
TA EMPREGO	oposição
TA FORÇA DE TRABALHO	associação implícita
TA NÍVEL DE DESEMPREGO	interfaceta
DESENVOLVIMENTO ECONÔMICO	
TA CRESCIMENTO ECONÔMICO	causalidade
TA DESEMPREGO	influência
TA POBREZA	influência
TA PROMOÇÃO INDUSTRIAL	instrumental

Fonte: Motta, 1987, p. 52.

As relações presentes nos tesouros são tratadas por relações lógicas, ontológicas e de efeito por Gomes (1990). Sendo que,

- a) Relacionamentos lógicos, oriundos da comparação de dois conceitos, podem se dividir em: genérico-específico; relacionamento analítico (relações associativas) e relacionamento de oposição;
- b) Relacionamentos ontológicos, reunindo as relações partitivas, relacionamentos de sucessão ou contigüidade e o relacionamento material-produto;
- c) Relacionamentos de efeito, constituído pelos relacionamentos de causalidade ou causa-efeito, instrumentalidade e os relacionamentos de descendência (relações genealógicas entre os termos).

Khoo e Na (2006), em capítulo de revisão sobre as relações semânticas na ciência da informação, sistematizam e apresentam as relações em paradigmáticas e sintagmáticas. Seriam as relações paradigmáticas: relação de Hiponímia-hiperonímia (equivalente ao conceito específico e geral); relação de Troponímia (a relação de hiponímia-hiperonímia entre dois verbos); relação Meronímia-Holonímia (a relação parte-todo, mas enquanto na Hiponímia, a relação é interna aos conceitos, na Meronímia, a relação dá-se entre conceitos); relação de partonímia (que pode se dividir em diferentes categorias de parte-todo: componente funcional/inteiro [roda de uma bicicleta]; segmentos do inteiro [o inteiro dividido em partes como uma torta]; Membros/Coleção [cão; matilha]; Subconjuntos de conjuntos [frutas; maçãs]; *social whole-staff* [aeronave; tripulação]; organização e seu chefe [navio; capitão]; inteiro e sua unidade uniforme [areia; grão]; inteiro/seu centro [montanha; pico da montanha])²⁰; relação de Sinonímia (senso-sinônimo: quando os termos compartilham um ou mais sentidos; quase-sinonímia: quando os termos compartilham sentidos diferentes, mas próximos no significado e a sinonímia-parcial: quando os termos compartilham sentidos, mas se diferem em algum aspecto) e a relação de Antonímia. Como relação sintagmática, os autores distinguem apenas a relação de causa e efeito.

A partir dos projetos e autores apresentados neste capítulo, apresentar-se-á os resultados da presente pesquisa, em capítulo próprio (Capítulo 4 - Resultados), com os demais resultados obtidos.

²⁰ No original: Functional component of a whole [wheel of a bicycle]; The segmented whole [the whole divided into pieces like a pie]; Members of a collection of elements / member-collection [dog; pack]; Subsets of sets [set inclusion, e.g. fruits and apples]; social whole-staff [aircraft; crew]; organization and its head [ship; captain]; a whole and its uniform unit [sand; grain]; a whole and its center or culmination [mountain; peak [of the mountain]] (KHOO; NA, 2006, p. 177).

3 METODOLOGIA

A proposta desta pesquisa é indicar as principais relações semânticas para subsidiar a representação de domínios do conhecimento através das ontologias. Esta pesquisa tem caráter exploratório e é de natureza qualitativa, pois privilegia a descrição das relações encontradas, para uso em pesquisas futuras em ontologias. Esta postura é justificada devido à limitação de tempo para o desenvolvimento da pesquisa, e principalmente por privilegiarmos o estudo da natureza das relações semânticas como forma de apoiar pesquisas futuras. Selltiz (1967, p. 60) destaca que,

estudos exploratórios podem ter outras funções que não a formulação de um problema para investigação mais exata ou para a criação de hipóteses. Estudos desta natureza objetivam aumentar o conhecimento do pesquisador acerca do fenômeno que deseja investigar e qualificar estudos futuros (SELLTIZ, 1967, p. 60).

A própria delimitação dos tipos de ontologias para composição de uma amostragem significativa de relações careceria de ser realizada após a revisão da literatura, porque apenas com este levantamento e a partir do estudo das relações existentes poder-se-ia fornecer parâmetros adequados para a caracterização das relações semânticas em ontologias.

Em síntese, os instrumentos empregados nesta pesquisa foram:

- a) A análise documental para sistematização das relações semânticas em ciência da informação e sistematização dos métodos para o estabelecimento das mesmas e;
- b) A análise conceitual a partir da teoria do conceito, que orientou o levantamento dos enunciados e consequentemente das características dos conceitos para a análise das definições (das relações), orientando a identificação dos seguintes tópicos sobre a relação semântica (termo): referente (objeto ou idéia do mundo) e as características (a soma dos enunciados verdadeiros levantados).

3.1 Análise documental: sistematização das relações semânticas em ciência da informação

Do levantamento bibliográfico, esperou-se obter a sistematização das relações semânticas no campo da OC e a sistematização dos métodos para o estabelecimento de relações em ontologias. A determinação deste levantamento está apresentada a seguir.

Para delimitar e sistematizar as relações em ciência da informação, a revisão de literatura foi orientada para identificar os trabalhos que se ocuparam minimamente em apresentá-las, com sua respectiva definição, nesta área de conhecimento. A literatura em Biblioteconomia apresenta inúmeros trabalhos sobre as relações entre os conceitos em linguagens e métodos de indexação, cuja teoria é muito influenciada por áreas como a Terminologia, a Lógica, a Lingüística. O esperado, considerando-se o crescente enriquecimento das relações semânticas em ferramentas para a OC, é que um levantamento histórico das relações semânticas em ferramentas tradicionais apontem, se não todas, as principais relações semânticas para apoiar o desenvolvimento de ontologias.

A busca por uma metodologia para o estabelecimento de relações, vai de encontro ao questionamento de Khoo e Na (2006, p. 187) ao afirmarem que

do que é de nosso conhecimento, nenhuma análise sistemática dos tipos de relações semânticas usadas em ontologias, foi reportada na literatura [...] Não há avaliação sistemática da eficiência em ontologias ou os vários tipos de relações semânticas que ocorrem em aplicações reais” (KHOO; NA, 2006, p. 187, *tradução nossa*).

Para atingir os objetivos deste estudo e propor um modelo de relacionamentos entre os conceitos em ontologias, fez-se necessário o levantamento bibliográfico retrospectivo – com objetivo de listar os tipos de relações existentes em biblioteconomia e ciência da informação. Este levantamento pretendeu ser exaustivo e privilegiou a inserção de trabalhos relevantes para a pesquisa. Para isso, a busca ocorreu a partir de manuais para construção e uso de instrumentos de representação e recuperação da informação; trabalhos de revisão; trabalhos que discutem a natureza das relações semânticas e os métodos de organização do conhecimento, além de trabalhos sobre suas bases teóricas e metodológicas. A literatura apontou as várias ferramentas desenvolvidas com o propósito de organização do conhecimento, dos quais foram selecionados os

seguintes projetos e ferramentas: o sistema unitermo; projeto Syntol; sistema PRECIS, indexação relacional; sistemas de classificação bibliográfica e tesouros. Interessante que a análise da literatura coincide com a própria evolução do desenvolvimento dos tesouros apontados em Lancaster (1986), em Campos (2001) e nos SOC listados na taxonomia proposta por Hodge (2000). Ferramentas estas que, somadas, indicam o panorama do tratamento às relações semânticas pela área. Seguem as referências que orientaram a temática da pesquisa:

1. AMERICAN NATIONAL STANDARDS ORGANIZATION. **ANSI/NISO Z 39.19: 2005**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005. 176 p. Disponível em: <<http://www.niso.org/standards/resources/Z39-19-005.pdf?CFID=2436046&CFTOKEN=81298864>>. Acesso em: 27 jun. 2007.
2. AUSTIN, D. The development of PRECIS: a theoretical and technical history. **Journal of documentation**, v. 30, n. 1, p. 97-102, 1974.
3. BEGHTEL, Clare. Semantic validity: concepts of warrant in bibliographic classification systems. **Library resources and technical services**, v. 30, n. 2, p. 109-25, 1986.
4. CAMPOS, Maria Luiza de Almeida. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: UFF, 2001.
5. CLARKE, S.G. D. Thesaural relationships. In C.A. BEAN; R. GREEN (Eds.). **Relationships in the organization of knowledge**. Dordrecht: Kluwer, 2001. p. 37-52.
6. CLASSIFICATION RESEARCH GROUP. The need for a faceted classification as the basis of all methods of information retrieval. UNESCO document 320/5515 (International advisory committee for documentation and terminology in pure and applied science). Paris, 1955 *apud* FROHMANN, B. P. An investigation of the semantic bases of some theoretical principles of classification proposed by Austin and the CRG. **Cataloging & Classification Quarterly**, v. 4, n. 1, 11-27. 1983.
7. COATES, E. J. Some properties in the structure of indexing languages. **Journal of documentation**, London, v. 29, n. 4, p. 390-404, dez. 1973.
8. CROS, R. C; GARDIN, J. C; LEVY, F. **L'Automatisation des recherches documentaires**: un modèle général, le Syntol. 2. ed. Paris: Gauthier-Villars, 1968.

9. CUNHA, Isabel M. R. Análise documentária. In: SMITH, J. (Coord.). **Análise documentária: análise da síntese**. Brasília: IBICT, 1987. cap. 4, p. 38-60.
10. DAHLBERG, I. A referent-oriented analytical concept theory of interconcept, **Intern.Classificat.**, Frankfurt, v. 5, n. 3, p. 142-150, 1978a.
11. _____. **Conceptual structures and sistematization**. International Forum on Information and documentation, v. 20, n. 3, p.09-24, Jul. 1995
12. _____. Futuro das linguagens de indexação. In: Conferencia Brasileira de Classificação Bibliográfica, 1976, Rio de Janeiro. **Anais...** Comemorativa do centenário da primeira edição de Classificação Decimal de Dewey. Rio de Janeiro: IBICT; Associação dos Bibliotecarios do Distrito Federal. 1979. p. 323-334. Disponível em: <<http://www.conexaorio.com/bit/dahlberg/index.htm>>. Acesso em: 28 jul. 2009.
13. _____. Knowledge organization: its scope and possibilities. **Knowledge Organization**, v. 20, n. 4, 211-222, 1993.
14. _____. Teoria do conceito. **Ciência da informação**. Rio de Janeiro, v. 7, n. 2, p. 101-107, jul./dez. 1978b.
15. FARRADANE, J. E. L. A scientific theory of classification and indexing and its practical application. **The Journal of Documentation**. v. 6, n. 2, p. 83-99, 1950.
16. _____. A scientific theory of classification and indexing: further considerations. **The journal of documentation**, v. 8, n. 2, p. 83-99, jun. 1952.
17. _____. Concept organization for information retrieval. **Information Storage and Retrieval**, 3(4), 297-314, 1967.
18. _____. Knowledge, information, and information science. **Journal of information science**, v. 2, p. 75-80, 1980a.
20. _____. Relational Indexing: part I. **Journal of information science**, n.1, p. 267-276, 1980b.
21. _____. Relational indexing: part II. **Journal of information science**, n. 1, p. 313- 324,

- 1980c.
22. _____. The psychology of classification. **Journal of Documentation**, v. 11, n. 4, p. 187-201, 1955.
 23. FOSKETT, A. C. **A abordagem temática da informação**. São Paulo: Polígono, 1973. 437p.
 24. FROHMANN, B. P. An investigation of the semantic bases of some theoretical principles of classification proposed by Austin and the CRG. **Cataloging & Classification Quarterly**, v. 4, n. 1, 11-27. 1983.
 25. GARDIN, Jean-Claude. Document analysis and linguistic theory. **Journal of documentation**, London, v. 29, n. 2. p. 137-168, jun. 1973.
 26. GHOSE, A. DHAWLE, A. S. Problems of thesaurus construction. **Journal of the American society for information science**, v. 28, n. 4, p. 211-217, 1977.
 27. GOMES, H. E. Biblioteconomia, Informação e Tecnologia da Informação. **Elaboração de tesouros documentários: tutorial**. Disponível em: <<http://www.conexaorio.com/bit/tesauro/index.htm>>. Acesso em: 29 jul. 2009.
 28. _____. (Org.). **Manual de elaboração de tesouros monolíngues**. Brasília: Programa Nacional de Bibliotecas de Instituições de Ensino Superior, 1990.
 29. HJORLAND, B. **Discussion of Dahlberg's theory of concepts and knowledge organization**. 2007a. Disponível em: <http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/discussion_of_dahlberg.htm>. Acesso em: 29 jul. 2009.
 30. _____. **Semantic relations: meaning relations**. 2007b. Disponível em: <http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/semantic_relations.htm>. Acesso em: 29 jul. 2009.
 31. _____. Semantics and Knowledge Organization. **Annual Review of Information Science and Technology**, v. 41, p. 367-405, 2007c.
 32. INTERNATIONAL STANDARD ORGANIZATION. **ISO 704: Terminology work:**

- Principles and methods principles and methods of terminology. Geneva: ISO, 2000.
33. KHOO, C. S. G.; NA, Jin-Cheon. **Semantic relations in Information Science**. Annual Review of Information Science and Technology, v. 40, p. 157-228, 2006.
 34. LANCASTER, F. W. **Vocabulary control for information retrieval**. Arlington, VA: Information Resources Press, 1986.
 35. _____. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 2004.
 36. MANIEZ, Jacques. Relationships in thesauri: some critical remarks. **International classification**, v. 15, n. 3, 1988. p. 133-138.
 37. MOTTA, Dilza Fonseca da. **Método relacional como nova abordagem para a construção de tesouros**. 1987. 89 f. Dissertação (Mestrado em ciência da informação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1987.
 38. PIEDADE, M. A. R. **Introdução à teoria da classificação**. 2. ed. Rio de Janeiro: Interciência, 1983. 221 p.
 39. RANGANATHAN, S. R. **Prolegomena to library classification**, Bombay: Ásia Publishing House, 1967. 640p *apud* CAMPOS, Maria Luiza de Almeida. *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói: Ed. da UFF, 2001.
 40. _____. **The Colon Classification**. New Brunswick, N.J.: Graduate School of Library Service, Rutgers University. 1965.
 41. RIVIER, Alexis. Construção de linguagens de indexação: aspectos teóricos. **Revista Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 21, n. 1, p. 56-99, jan./jun. 1992.
 42. SVENONIUS, E. The intellectual foundation of information organization. Cambridge, MA: MIT Press, 2000.
 43. VALE, E. A. do. Linguagens de indexação. In: SMITH, J. (Coord.). **Análise documentária: análise da síntese**. Brasília: IBICT, 1987. cap. 1, p. 12-26. p. 20.
 44. WILLETS, Margaret. An investigation of the nature of the relation between terms in

thesauri. **Journal of Documentation**, v. 31, n. 3, p. 158-184, 1975.

45. WUESTER, E. L'étude scientifique générale de la terminologie, zone frontalière entre la Linguistique, la Logique, l'Ontologie, l'Informatique et les Sciences des Choses. In: RONDEAU, G.; FELBER, F. (Org.). Textes choisis de terminologie: fondements théoriques de la terminologie. Québec: GIRSTERM, 1981. p. 57-114 *apud* RIVIER, Alexis. Construção de linguagens de indexação: aspectos teóricos. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 21, n. 1, p. 56-99, jan./jun. 1992.
46. YU, C. T.; RAGHAVAN, V. V. Single-pass method for determining the semantic relationships between terms. **Journal of the American Society for Information Science**. V. 28, n. 6, p. 345-354, 1977.

A forma como ocorreu a aplicação da análise conceitual - a partir da análise documental - está explicitada a seguir.

3.2 Análise conceitual com a aplicação da teoria do conceito: enunciados verdadeiros para as relações semânticas

Por análise conceitual compreende-se o processo intelectual que, associado a outras técnicas, envolve a identificação, a determinação, a estruturação e a validação dos conceitos e relações semânticas apropriados para a representação do conhecimento, que em SOC permite ações para a modelagem conceitual.

Nesta pesquisa, optou-se por adotar a teoria do conceito (DAHLBERG, 1978b) como método para a análise conceitual ao agrupar as definições para as relações semânticas, visto que permite o levantamento dos enunciados verdadeiros sobre cada relação. Acredita-se que, através da análise e comparação das ‘intensões’ dos conceitos das relações semânticas apontados na literatura, chegar-se-á a um núcleo de relações essenciais úteis para a construção de ontologias.

Por intensão, adotou-se a definição de Dahlberg (1978b) que indica:

A intensão do conceito é a soma total das suas características. É também a soma total dos respectivos conceitos genéricos e as diferenças específicas ou características especificadoras. Na representação da intensão do conceito numa definição nem todos os conceitos genéricos necessitam ser mencionados. É suficiente mencionar o mais próximo possível já que este necessariamente contém os demais [...] Já a extensão do conceito 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 (DAHLBERG, 1978b, p. 105).

Foi em Dahlberg (1978a; 1978b) que o “conceito de conceito” foi seguramente definido. Com inúmeros trabalhos na área e com a teoria do conceito (TC), o seu trabalho apresenta o conceito como elemento para refletir a própria teoria da OC. Para esta autora, o conceito é formado por três elementos: o referente (aquilo que se pretende conceituar), as características (soma dos enunciados verdadeiros sobre o referente) e a forma verbal – formando o triângulo do conceito. Deste modo, o conceito passa de um artefato mental, abstrato, para instrumentar sua pesquisa e sua manipulação para a construção de sistemas terminológicos. Diferente do Triângulo do Significado, de Ogden e Richards (1972, p. 32), onde o conceito é um dos vértices do triângulo (e

o seu significado pode assumir várias possibilidades), a tríade proposta por Dahlberg apresenta o conceito como a soma dos enunciados verdadeiros/essenciais sobre um referente e o termo/léxico é a forma comunicável e representável do conceito.

Ainda de acordo com Dahlberg (1993), o conceito é a “unidade do conhecimento”. Hjørland (2007a) recorda que, por compreender um enorme campo de disciplinas, a aplicação da TC deve resultar em análises ampliadas dos conceitos.

Este método foi também utilizado em Motta (1987), para distinguir tipos diferenciados de relações associativas em tesouros:

Para se usar conceitos e termos correta e precisamente, é necessário saber o que os conceitos representam, qual é sua intensão, isto é, qual é a soma de suas características. Isso pode ser conseguido definindo-se o conceito, já que a definição estabelece os limites de sua intensão (MOTTA, 1987, p. 39). [Para isso, a autora propõe o método relacional] Método que pretende ser instrumento para o estabelecimento de relações em tesouros de forma mais objetiva. O método pode ser definido da seguinte forma: método analítico que consiste na análise das definições dos conceitos que integrarão o sistema, de forma a identificar suas características, e na estruturação desses conceitos, tendo em vista as relações existentes entre eles. (MOTTA, 1987, p.39)

Na teoria do conceito (DAHLBERG, 1978a; 1978b), a autora verifica os fatores que determinam e interferem na formação e definição dos conceitos. Através da teoria do conceito, podemos expressar o referente (A), as características (B) - as asserções corretas a cerca do item, os enunciados representam um atributo de um objeto (característica) e a forma verbal (C), como ilustrado a seguir (FIG. 10):

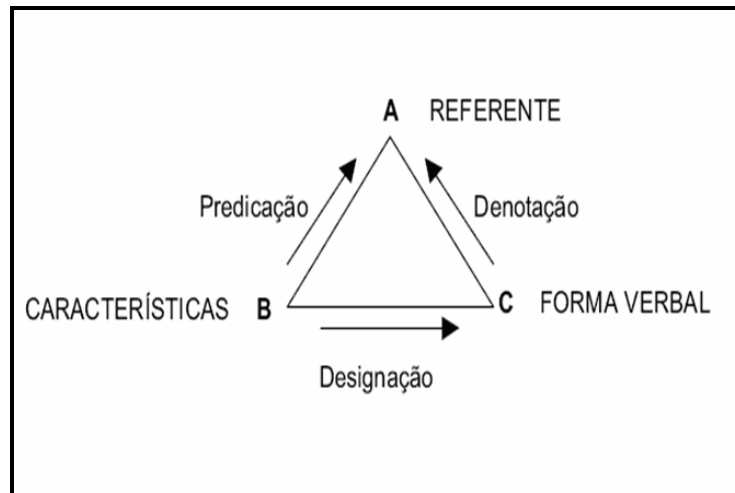


Figura 10 - Triângulo do conceito

Fonte: CAMPOS, 2001, p. 103.

A teoria do conceito aplicada a esta pesquisa não objetiva a definição do conceito (da relação semântica analisada), mas sim levantar as características que constituem a relação, a partir das definições presentes na literatura. A partir das citações de autores, exclusivamente da área, agrupar-se-á as características que tornam unívocas cada relação (somadas, as características fornecem uma definição mais apropriada aos conceitos).

Um exemplo da aplicação deste método está ilustrado a seguir (QUADRO 06):

Quadro 06 - Eixos da teoria do conceito para o estudo das relações semânticas

FORMA VERBAL	RELAÇÃO GENÉRICA
Citações para a relação na literatura	<p>A relação gênero-espécie, também conhecida como relação de inclusão, é uma relação hierárquica clássica com as propriedades de reflexibilidade, transitividade, e assimetria. Ela tem outra propriedade, a qual na literatura computacional é chamada de herança e na literatura de classificação <i>força hierárquica</i>, de forma que o que é uma verdade de uma categoria dada (móvel) é verdade de todas as categorias que se agrupam (SVENONIUS, 2000, p.163).</p> <p>Uma relação genérica ocorre entre dois conceitos quando a intensão do conceito subordinado (conceito específico) inclui a intensão do conceito superordenado (conceito genérico) mais pelo menos uma característica delimitante adicional. O conceito superordenado em uma relação genérica é chamado de conceito genérico e o conceito subordinado é chamado conceito específico. Em uma relação genérica, há uma relação inversa entre a intensão de um conceito e sua extensão. (ISO 704, 2000, p. 5)</p>
Referentes	Categoria (móvel) categorias que ela classifica (cadeiras, mesas, etc.) (SVENONIUS, 2000, p. 163).

Fonte: Desenvolvido pela autora.

A partir das citações, podemos enunciar as seguintes características sobre a relação genérica:

- 1- É uma relação de inclusão.
- 2- É uma relação hierárquica clássica.
- 3- É uma relação reflexiva.
- 4- É uma relação transitiva.
- 5- É uma relação assimétrica.
- 7- É uma relação lógica.
- 8- Possui propriedade de herança: o que é verdade para o termo superordenado (conceito genérico), aplica-se ao termo subordinado (conceito específico).
- 9- A intensão do conceito específico inclui a intensão do conceito genérico, mais pelo menos uma característica delimitante adicional.

Desta forma, acredita-se que ficariam evidenciados os limites para a aplicação da relação semântica (enquanto conceito), o que é propósito no presente trabalho. A síntese dos tipos de relações semânticas existentes e os resultados da aplicação desse método podem ser conferidos no capítulo a seguir.

4 RESULTADOS

A presente pesquisa pretendeu indicar a possibilidade de identificação de um núcleo básico de relações semânticas para o desenvolvimento de ontologias. De acordo com a literatura sobre trabalhos anteriores que envolvem a temática das relações semânticas para a construção de sistemas terminológicos (detalhados no referencial teórico, capítulo 2), chegou-se aos seguintes trabalhos que apresentam, originalmente, intensões para as relações semânticas:

1. AMERICAN NATIONAL STANDARDS ORGANIZATION. ANSI/NISO Z39.19:2005: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005. 176 p. Disponível em: <<http://www.niso.org/standards/resources/Z39-19-005.pdf?CFID=2436046&CFTOKEN=81298864>>. Acesso em: 27 jun. 2007.
2. DAHLBERG, I. Teoria do conceito. *Ciência da informação*. Rio de Janeiro, v. 7, n. 2, p. 101-107, jul./dez. 1978b.
3. FARRADANE, J. E. L. Relational Indexing: part I. *Journal of information science*, n.1, p. 267-276, 1980b.
4. HJORLAND, B. Semantics and Knowledge Organization. *Annual Review of Information Science and Technology*, v. 41, p. 367-405, 2007c
5. INTERNATIONAL STANDARD ORGANIZATION. ISO 704: Terminology work: Principles and methods principles and methods of terminology. Geneva: ISO, 2000.
6. KHOO, C. S. G.; NA, Jin-Cheon. Semantic relations in Information Science. *Annual Review of Information Science and Technology*, v. 40, p. 157-228, 2006.
7. MANIEZ, Jacques. Relationships in thesauri: some critical remarks. *International classification*, v. 15, n. 3, 1988. p. 133-138.
8. MOTTA, Dilza Fonseca da. Método relacional como nova abordagem para a construção de tesouros. 1987. 89 f. Dissertação (Mestrado em ciência da informação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1987.

9. SVENONIUS, E. The intellectual foundation of information organization. Cambridge, MA: MIT Press, 2000.

Para a presente pesquisa, podem-se indicar os seguintes resultados:

- a) síntese das relações semânticas (agrupamento sugerido pela presente pesquisa) indicadas na literatura (QUADRO 08), cuja notação para as relações será adotada nos demais quadros.
- b) síntese das relações semânticas existentes, indicadas por autor de origem (QUADRO 09). O quadro 09 mapeia a origem das relações resultantes da presente pesquisa e,
- c) síntese dos enunciados, com a apresentação das principais características indicadas pela análise conceitual das relações semânticas existentes (Seção 4.1).

A análise conceitual sobre as relações semânticas permitiu que as mesmas fossem agrupadas da forma indicada no quadro 08.

Quadro 07 – Síntese de relações semânticas

A	EQUIVALÊNCIA	
A1		Sinonímia total
A2		Variações lexicais
A3		Quase-sinonímia ou Sinonímia-parcial
A4		Antonímia
B	HIERÁRQUICA	
B1		Gênero-espécie
B2		Instanciação
B3		Troponímia
C	PARTITIVA	
C1		Componente funcional/inteiro
C2		Segmentos do inteiro
C3		Membros/coleção
C4		Subconjuntos de conjuntos
C5		<i>Social whole/staff</i>
C6		Organização/seu chefe
C7		Inteiro/sua unidade uniforme
C8		Inteiro/seu centro
C9		Ação/propriedade
C10		Conceito ou objeto/propriedade
C11		Material/propriedade
C12		Matéria ou substância/propriedade
D	CAUSA - EFEITO	
D1		Ação/produto
D2		Ação/propriedade
D3		Ação/alvo
D4		Ação/equipamento-ferramenta
D5		Ação/lugar-localização
D6		Ação/ator
E	ASSOCIATIVA	
E1		Processo/agente
E2		Processo/contra-agente
		(Continua...)

	(Conclusão)
E3	Conceito ou objeto/origem
E4	Conceito ou objeto/unidade ou mecanismo de medida
E5	Matéria bruta/produto
E6	Disciplina ou campo/objeto ou profissional
E7	Matéria prima/produto
E8	Quantidade/unidade
E9	Material/estado
E10	Item concreto/material
E11	Item concreto/forma
E12	Recipiente/o que é contido
E13	Atividade/ferramenta
E14	Passos de um ciclo
E15	Produtor/produto
E16	Duração/equipamento de medida
E17	Profissão/ferramenta típica
E18	Objeto/ferramenta associada
E19	Organização/edificação associada

Fonte: Desenvolvido pela autora.

Desta primeira análise, foram excluídas as seguintes tipologias para as relações semânticas: Relações funcionais (DAHLBERG, 1978b); relação de concomitância ou coincidência (FARRADANE, 1980b) e a relação interfaceta (MOTTA, 1987), por apresentarem caráter ambíguo e não apresentarem definição similar em outros autores.

Foram excluídos ainda, os conceitos de paradigmático e sintagmático (CROS; GARDIN; LEVY, 1968); lógico e ontológico (WUESTER, 1981 apud CAMPOS, 2001 e GOMES, 1990) e temporal/permanente (FARRADANE, 1980b), mas os mesmos poderão ser retomados na síntese dos enunciados dos tipos de relações, quando se tratarem de características essenciais para as relações. As citações estão identificadas na seção a seguir, seguida da síntese destes elementos para os conceitos propostos.

Quadro 08 - Síntese das relações semânticas por autor de origem

RELAÇÃO SEMÂNTICA		AUTORES					
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987)/ Tesauros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
A - EQUIVALÊNCIA							
A1	Sinonímia total	..	SIM	SIM	..	SIM	SIM
A2	Variações lexicais	SIM	..
A3	Quase-sinonímia ou Sinonímia-parcial	SIM	SIM
A4	Antonímia	SIM	SIM	SIM	..	SIM	SIM
B - HIERÁRQUICA							
B1	Gênero-espécie	SIM	SIM	SIM	SIM	SIM	SIM
B2	Instanciação	SIM	..
B3	Troponímia	SIM
C - PARTITIVA		SIM	SIM	SIM	SIM	SIM	SIM
C1	Componente funtional/inteiro	SIM
C2	Segmentos do inteiro	SIM
C3	Membros/coleção	SIM
C4	Subconjuntos de conjuntos	SIM
C5	<i>Social whole/staff</i>	SIM
C6	Organização/seu chefe	SIM
C7	Inteiro/sua unidade uniforme	SIM
C8	Inteiro/seu centro	SIM
(Continua...)							

RELAÇÃO SEMÂNTICA		AUTORES					(Continuação)
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987)/ Tesauros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
D - CAUSA-EFEITO		..	SIM	SIM	SIM	SIM	SIM
E - ASSOCIATIVA		..	SIM	SIM	SIM	SIM	..
E1	Processo/agente	SIM	..
E2	Processo/contra-agente	SIM	..
E3	Conceito ou objeto/origem	SIM	..
E4	Conceito ou objeto/ unidade ou mecanismo de medida	SIM	..
E5	Matéria bruta/produto	SIM	..
E6	Disciplina ou campo/objeto ou profissional	SIM	..
E7	Matéria prima/produto	SIM
D4	Ação-equipamento/ferramenta	SIM
E8	Quantidade-unidade	SIM
E	Material/propriedade	SIM
E9	Material/estado	SIM
C1	Matéria ou substância-propriedade	SIM
E10	Item concreto/material	SIM
E11	Item concreto/forma	SIM
D5	Ação/lugar-localização	SIM
D6	Ação/ator	SIM
							(Continua...)

RELAÇÃO SEMÂNTICA		AUTORES					(Conclusão)
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987)/ Tesauros	ISO 704 (2000)	ANSI (2005)	
E12	Recipiente/o que é contido	SIM
E13	Atividade/ferramenta	SIM
E14	Passos de um ciclo	SIM
E15	Produtor/produto	SIM
E16	Duração/equipamento de medida	SIM
E17	Profissão/ferramenta típica	SIM
E18	Objeto/ferramenta associada	SIM
E19	Organização/edificação associada	SIM

4.1 Enunciados e características das relações semânticas

A partir das citações identificadas na literatura sobre os tipos de relações semânticas, foram identificados os enunciados verdadeiros para cada uma das relações semânticas, seguido dos referentes (objeto ou idéia do mundo), caso apresentado pelo autor de origem, conforme a tríade proposta por Dahlberg (1978a; 1978b). Ao final de cada conjunto de citações, apresentar-se-á a síntese dos elementos do conceito, identificados pela pesquisa.

A - RELAÇÃO DE EQUIVALÊNCIA

A1 – Sinonímia Total

Dahlberg (1978b) Relação quantitativa de identidade: “As características do conceito A (x,x,x) e do conceito B (x,x,x) são as mesmas” (DAHLBERG, 1978b, p. 104).

Referentes: Não apresenta.

Farradane (1980b) Quando dois conceitos têm muito em comum, cuja percepção é simples, sem referência ao tempo. Dá-se pela expressão de algum grau de equivalência ou equivalência completa entre conceitos (como o caso do sinônimo). Trata-se de uma relação permanente e indistinta (FARRADANE, 1980b, p. 271).

Referentes: Sódio/=íon
Melaços de cana/= forragem
Acetona /= solvente

ANSI (2005) A relação de equivalência é aquela que representa os sinônimos ou quase sinônimos de um termo. Divide-se em: relações de sinonímia; as variações lexicais entre conceitos e a quase-sinonímia. Aparecem com a indicação de USE ou UP (use para), nos tesouros, mas não há distinção sintática para sinônimos e quase-sinônimos (ANSI, 2005, p. 43).

Referentes: UN/*United Nations* (ANSI, 2005, p. 42)

Khoo e Na (2006) “Sinônimos são geralmente tratados como reflexivos, simétricos e

transitivos” (KHOO, NA, 2006, p. 178).

Referentes:

Não apresenta.

Maniez (1988)

Este relacionamento [equivalência] oferece proteção essencial contra o silêncio [no momento da recuperação da informação], uma vez que, implica em que: 'se você está interessado em A, então você está intrinsecamente interessado em B (MANIEZ, 1988, p. 136).

"Equivalência estrita dos termos não pode existir" (MANIEZ, 1988, p. 136).

Referentes:

Não apresenta.

Svenonius (2000)

Relações entre duas ou mais variações ortográficas, variantes sintáticas ou sinônimos são relações de equivalência. Relacionamentos de Equivalência têm propriedades de simetria, reflexividade e transitividade. [...] Eles [os termos] são “inter-substituíveis” (*intersubstitutable*) em (quase) todos os contextos, inclusive na recuperação da informação, onde a pesquisar por um dos termos sinônimos poderá recuperar todos os documentos que poderiam ser obtidas através de pesquisas com o outro termo sinônimo (SVENONIUS, 2000, p. 156).

Referentes:

Não apresenta.

Hjorland (2007c)

“A denota o mesmo que B; A é equivalente a B” (HJORLAND, 2007c, p. 405).

Referentes:

Não apresenta.

Síntese dos enunciados – elementos do conceito “Sinonímia Total”

1. *É uma relação paradigmática.*
2. *É uma relação lógica.*
3. *É atemporal, permanente.*
4. *Os conceitos A e B apresentam as mesmas características: A denota o mesmo que B.*
5. *Os conceitos apresentam equivalência completa.*
6. *Os conceitos A e B são indistintos.*

7. *Os conceitos A e B podem ser substituídos em vários domínios.*
8. *Apresenta propriedade de simetria (A relação é a mesma, não importa a direção dos conceitos [A = B e B = A]).*
9. *É uma relação reflexiva.*
10. *É uma relação transitiva (A = B = C, então A = C).*

A2 - Variações Lexicais

ANSI (2005)

Varição lexical difere dos sinônimos porque sinônimos são diferentes termos para o mesmo conceito, enquanto a variação lexical são diferentes formas da palavra para a mesma expressão. A variação pode ser recorrente da ortografia ou variação gramatical ou formatos abreviados (ANSI, 2005, p. 45).

Referentes:

Pediatrics; paediatrics (pediatria)

Síntese dos enunciados – elementos do conceito “Variações Lexicais”

1. É uma relação paradigmática.
2. É uma relação lógica.
3. É atemporal, permanente.
4. Os conceitos A e B apresentam as mesmas características: A denota o mesmo que B.
5. – Os conceitos A e B equivalem ao mesmo conceito.
6. - Os conceitos A e B diferem-se quanto à ortografia.
7. – Variação lexical é um tipo de relação de equivalência.

A3 - Quase-Sinônimos / Sinonímia Parcial

Dahlberg (1978b) Relação quantitativa de intersecção: “As características do conceito A (x,x,o) coincidem em algum elemento com as características do conceito B (x,o,o)” (DAHLBERG, 1978b, p.104).

Referentes: Não apresenta.

ANSI (2005) “Um termo cujo significado não é exatamente sinônimo com o de outro termo, ainda o qual pode, todavia, ser tratado como seu equivalente em um vocabulário controlado” (ANSI, 2005, p. 163).

Referentes: Água do mar / água salgada;
Doçura / aspereza (antônimos).

Khoo e Na (2006) Tipos comuns de sinonímia são: sense-synonyms (termos que compartilham um ou mais sentidos), Sinônimos parciais [near synonyms] (os quais não tem sentidos idênticos, mas são parecidos em significados), e sinônimos parciais [partial synonyms] (os quais dividem alguns sentidos, mas diferem em algum aspecto, por exemplo, na forma que eles são usados ou em alguma dimensão de significado) (CRUSE, 1986; LYONS, 1995 apud KHOO e NA, 2007, p. 177).

Referentes: Não apresenta.

Síntese dos enunciados – elementos do conceito “Quase-Sinônimos / Sinonímia Parcial”

1. - É uma relação lógica.
2. - É uma relação paradigmática.
3. - É atemporal, permanente.
4. - É um tipo de relação de equivalência.
5. - Os conceitos A e B têm significados aproximados.
6. Algumas características do conceito A são as mesmas características do conceito B, ou seja, há intersecção de características.
7. Quase-sinônimo ou sinônimo parcial são um tipo de relação de equivalência.

A4 – Antonímia

Dahlberg (1978b) Relação de oposição: podem ser das seguintes espécies: contradição e contrariedade. [...] Relação de oposição se aplicam principalmente a conceitos que expressam propriedades (DAHLBERG, 1978b, p. 105).

Relação de Disjunção: “Nenhuma característica em comum em A (x,x,x) e B (o,o,o)” (DAHLBERG, 1978b, p. 104).

Referentes: Contradição: numérico e não numérico; presente e ausente.
Contrariedade: preto e branco.

Farradane (1980b) Relação de distinção (*distinctness*): quando dois conceitos podem ser completamente discriminados, cuja percepção é simples, sem referência ao tempo. Aplica-se para expressar relação de imitação ou substituição. É atemporal, mas ao contrário da relação de co-ocorrência, é uma relação que expressa a distinção entre os conceitos (FARRADANE, 1980b, p. 270).

Referentes: homem /) estátua; recuperação da informação /) modelo matemático.

Motta (1987) Relação de oposição (MOTTA, 1987, p. 48).

Referentes: Emprego TR Desemprego

ANSI (2005) “Antônimos também podem ser tratados como termos relacionados, em vez de termos equivalentes [Quase-sinônimos]” (ANSI, 2005, p. 45).

Referentes: Baixo RT alto (ANSI, 2005, p. 56).

Khoo e Na (2006) O antônimo é não-reflexivo, simétrico e intransitivo (KHOO, NA, 2006, p. 178).

Referentes: *Quente/frio; sensual/frígido; grande/pequeno; bom/ruim, bom/mal [No original, hot/cold; steamy/frigid; big/small; good/bad, good/evil].*

Síntese dos enunciados – elementos do conceito “Antonímia”

1. É uma relação lógica.
2. É uma relação paradigmática.
3. É uma relação atemporal.
4. É uma relação não-reflexiva.
5. É uma relação intransitiva.
6. É uma relação simétrica.
7. Os conceitos A e B são distintos.
8. Os conceitos A e B apresentam características contrárias; opostas.

B - RELAÇÃO HIERÁRQUICA

Dahlberg (1978b) “Relações hierárquicas: quando dois conceitos diferentes possuem características, sendo que um deles possui uma característica a mais que o outro” (DAHLBERG, 1978b, p. 104).

Referentes: Árvore; Árvore frutífera; macieira (Relações hierárquicas); Árvore frutífera (macieira; pereira; pessegueiro); árvore de nozes (amendoeira; aveleira; nogueira).

Farradane (1980b) Relação de pertença (*appurtenance*): quando os conceitos têm muito em comum e cuja associação entre eles é fixa, permanente. Expressa relação de todo-parte, de gênero-espécie e toda propriedade física intrínseca de um material ou coisa particular. Ex.: mesa /(perna; gênero /(espécie; chá /(cafeína. Expressa ainda, relação para todas as propriedades físicas e intrínsecas de um objeto em particular [por exemplo: metal /(densidade; tubo /(diâmetro] (FARRADANE; 1980b, p.271).

Referentes: Mesa /(perna; gênero /(espécie; chá /(cafeína; Metal /(densidade; tubo /(diâmetro).

ISO 704 (2000) “A norma ISO 704 divide as relações hierárquica em relação genérica e relação partitiva” (ISO 704, 2000, p. 5).

Referentes:**Não apresenta.**

ANSI (2005)

“A relação hierárquica engloba: relação do tipo genérica (gênero-espécie); a relação de instanciação a relação do tipo partitiva (todo-parte)” (ANSI, 2005, p. 47).

Relações hierárquicas se baseiam em níveis de superordenação e subordinação, onde o termo superordenado representa uma classe ou um inteiro, e os termos subordinados referem aos seus membros ou partes (ANSI, 2005, p. 114).

Relação hierárquica: “Uma relação entre dois ou entre vários termos em um vocabulário controlado que descreve mais amplamente (genérico) para mais estritamente (específico) ou relações todo-parte” (ANSI, 2005, p. 160).

Referentes:

Vertebrados – mamíferos (ANSI, 2005, p. 46).

Maniez (1988)

Relações hierárquicas: por razões pragmáticas, este atributo muito vago aplica-se de modo geral aos gêneros/espécies e as relações todo/parte, as quais são com frequências misturadas em tesouros. Em princípio, estas relações são claramente distintas: um polegar é um dedo (gênero/espécie), mas não é uma mão (todo/parte). Segue que uma afirmação que seja verdadeira para um gênero é necessariamente verdade de uma parte. “Se você está interessado em dedos então você está certamente interessado em polegares – ainda que em uma extensão menor – mas “se você está interessado” em mãos então você provavelmente está interessado em dedos (grifo nosso) [...] A relação de gêneros/espécies é a base da classificação e silogismo. Ela é assimétrica e transitiva, uma vez que ela cria uma hierarquia perfeita (MANIEZ, 1988, p. 136, grifo nosso).

Referentes:**Não apresenta.****Síntese dos enunciados – elementos do conceito “relação hierárquica”**

1. Dois conceitos diferentes A e B, sendo que um deles possui uma característica a mais que o outro.
2. Relações hierárquicas se baseiam em níveis de superordenação e subordinação.
3. Os conceitos A e B são distintos.
4. Relação hierárquica pode expressar relação de gênero-espécie; relação todo-parte e relação de instanciação.

B1 - Relação Genérica

Dahlberg (1978b) Inclusão: As características do conceito A (x,x) estão contidas no conceito B (x,x,x) (DAHLBERG, 1978b, p. 104)

Referentes: Não apresenta.

Farradane (1980b) Relação de pertença (*appurtenance*): quando os conceitos têm muito em comum e cuja associação entre eles é fixa, permanente. Expressa relação de todo-parte, de gênero-espécie e toda propriedade física intrínseca de um material ou coisa particular. Ex.: mesa /(perna; gênero /(espécie; chá /(cafeína. Expressa ainda, relação para todas as propriedades físicas e intrínsecas de um objeto em particular (por exemplo: metal /(densidade; tubo /(diâmetro) (FARRADANE, 1980b, p.271).

Referentes: Mesa /(perna; gênero /(espécie; chá /(cafeína.
Metal /(densidade; tubo /(diâmetro).

ISO 704 (2000, p. 5) Uma relação genérica ocorre entre dois conceitos quando a intensão do conceito subordinado (conceito específico) inclui a intensão do conceito superordenado (conceito genérico) mais pelo menos uma característica delimitante adicional. O conceito superordenado em uma relação genérica é chamado de conceito genérico e o conceito subordinado é chamado conceito específico (ISO 704, 2000, p. 5).

Referentes: Instrumento de escrita (lápiz; caneta, marcador de texto);
Lápiz (lápiz [*lead pencil*]; lapiseira).

ANSI (2005) Esta relação [genérica] identifica a ligação entre uma classe e seus membros ou espécies. Este tipo de relação é frequentemente chamado de é um <IsA>. Uma forma simples de validar isso é formular a sentença lógica todos-alguns. Por ex.: plantas suculentas e cactos - alguns membros da classe plantas suculentas são conhecidas como cactos e todos os cactos, por definição, independentemente do contexto, são plantas suculentas (ANSI, 2005, p. 47).

Referentes: Plantas suculentas; cactos [alguns membros da classe “plantas suculentas” são os cactos enquanto todos os cactos são, por definição e independente do contexto, plantas suculentas (ANSI, 2005, p. 47)].

Khoo e Na (2006) Relação de Hiponímia-hiperonímia - A relação hipônima tem sido

aludida na literatura sob nomes diversos, incluindo ISA (é-um), um-tipo-de, taxonômica, superior-subordinado, gênero-espécies e relações classe-subclasse. Hiponímia refere-se ao termo/conceito mais específico (por exemplo, cão pastor alsaciano), e hiperonímia é o termo/conceito mais genérico (cachorro). A relação implica inclusão de classe (KHOO, NA, 2006, p. 174).

Referentes:

Cachorro; cão pastor alsaciano.

Svenonius (2000)

A relação gênero-espécie, também conhecida como relação de inclusão, é uma relação hierárquica clássica com as propriedades de reflexividade, transitividade e assimetria. Ela tem uma outra propriedade que na literatura computacional é chamada de herança e na literatura de classificação força hierárquica, por meio do qual o que é verdade de uma categoria dada (móvel) é verdade de todas as categorias que se agrupam (SVENONIUS, 2000, p.163).

Referentes:

Categoria (móvel) categorias que ela classifica (cadeiras, mesas, etc.).

Síntese dos enunciados – elementos do conceito “relação genérica”

1. É uma relação lógica.
2. É uma relação paradigmática.
3. É uma relação atemporal, permanente.
4. É uma relação reflexiva.
5. É uma relação transitiva.
6. É uma relação assimétrica.
7. Os conceitos A e B são distintos.
8. O conceito B (subordinado) possui uma característica a mais que o conceito A (superordenado).
9. As características do conceito A (x,x) estão contidas no conceito B (x,x,x).
10. A intensão do conceito específico inclui a intensão do conceito genérico mais pelo menos uma característica delimitante adicional.
11. A relação genérica pode ser chamada de relação de

inclusão, é-um, 'IsA', um-tipo-de, taxonômica, superior-subordinado, gênero-espécies e relações classe-subclasse.

12. A relação genérica identifica a ligação entre uma classe e seus membros ou espécies.
13. Em uma relação genérica, alguns membros de uma determinada classe C são espécies x, enquanto todas as espécies x são, por definição e independente do contexto, pertencentes a classe C.
14. A relação genérica é uma relação hierárquica clássica com as propriedades de reflexividade, transitividade e assimetria.
15. A relação genérica tem propriedade chamada de herança ou força hierárquica, ou seja, o que é verdade de uma classe é verdade para todas as espécies que se agrupam.

B2 - Relação de instanciação

DAHLBERG (1978b)

Toda vez que o objeto é pensado como único, distinto dos demais, constituindo uma unidade inconfundível (coisas, fenômenos, processos, acontecimentos, atributos, etc.) pode-se falar de *objetos individuais*. Pode-se dizer que o que caracteriza os objetos individuais é a presença das formas do tempo e espaço [...]. Mas, além dos objetos individuais, expressos pelos conceitos individuais, podemos referir-nos a *objetos gerais* que, de certo modo, prescindem das formas do tempo e espaço. A esses objetos situados fora do tempo e do espaço, correspondem os chamados conceitos gerais (DAHLBERG, 1978b, p. 101-102, grifo nosso).

Referentes:

IBICT (Instituto Brasileiro de Informações em Ciência e Tecnologia); Instituição.

ANSI (2005)

Este relacionamento identifica a ligação entre uma categoria geral

de coisas ou eventos, expressos por um nome comum e uma instância individual desta categoria, frequentemente, um nome próprio. Relações deste tipo são conhecidas como relações “IsA”(ANSI, 2005, p. 48).

Referentes:

Regiões montanhosas; Alpes; Himalaias (Alpes e Himalaias ocupam posições subordinadas em uma hierarquia – não são nem tipos, nem parte de regiões montanhosas, mas representam exemplos específicos ou instâncias) (ANSI, 2005, p. 48).

Síntese dos enunciados – elementos do conceito “relação de instanciação”

1. É uma relação lógica.
2. É uma relação paradigmática.
3. É uma relação atemporal, permanente.
4. Os conceitos A e B são distintos.
5. O conceito A refere-se a um conceito geral enquanto o conceito B (instância) refere-se a um conceito individual.
6. As características do conceito A (x,x) estão contidas no conceito B (x,x,x).
7. A relação genérica tem propriedade chamada de herança ou força hierárquica, ou seja, o que é verdade de uma classe é verdade para todas as espécies que se agrupam.

Khoo e Na (2006, p. 175) Troponímia refere-se à relação de hponímia-hiperonímia, mas quando da ocorrência entre verbos.

Referentes: Cantarolar (*Yodel*)/cantar; murmurar / falar.

Síntese dos enunciados – elementos do conceito “relação de troponímia”

1. Os conceitos A e B são verbos.
2. O conceito A é mais genérico que o conceito B.
3. O conceito B é mais específico que o conceito A.

C - RELAÇÃO PARTITIVA

Dahlberg (1978b) A relação partitiva existe entre um todo e suas partes. [...] Constitui também relação partitiva a que existe entre um produto e os elementos que o constituem (DAHLBERG, 1978b, p. 104).

Relações partitivas aplicam-se principalmente a conceitos que expressam *objetos* (DAHLBERG, 1978b, p. 105).

Referentes: Árvore (raízes, tronco, galhos, folhas, flores, frutos).

Farradane (1980b) Relação de pertença (*appurtenance*): quando os conceitos têm muito em comum e cuja associação entre eles é fixa, permanente. Expressa relação de todo-parte, de gênero-espécie e toda propriedade física intrínseca de um material ou coisa particular. Ex.: mesa /(perna; gênero /(espécie; chá /(cafeína. Expressa ainda, relação para todas as propriedades físicas e intrínsecas de um objeto em particular (por exemplo: metal /(densidade; tubo /(diâmetro) (FARRADANE, 1980b, p. 271).

Referentes: mesa /(perna; gênero /(espécie; chá /(cafeína.
metal /(densidade; tubo /(diâmetro).

ISO 704 (2000) Uma relação partitiva ocorre quando o conceito superordenado representa um inteiro, enquanto o conceito subordinado representa

parte deste inteiro. As partes, juntas, formam um inteiro. O conceito superordenado em uma relação partitiva é chamado o conceito compreensivo (*comprehensive concept*) e o conceito subordinado é chamado conceito partitivo. Conceitos subordinados em um mesmo nível, que compartilham a mesma dimensão são chamados conceitos coordenados (ISO 704, 2000, p. 8).

Conceitos subordinados que estão no mesmo nível e dividem a mesma dimensão são chamados conceitos coordenados. Relações partitivas, como a relações genéricas podem ser expressas em séries verticais e horizontais. As partes que formam um inteiro podem ser similares em natureza (átomo em uma molécula de oxigênio) ou distintas. Uma ou mais partes podem ser obrigatórias (essenciais) ou opcional (não-essencial) [...] (ISO 704, 2000, p. 8).

Referentes:

Variável dentro de um domínio (por exemplo, uma caneta pode ter como parte: um reservatório [*reservoir*], um cartucho [*cartridge*] ou um refil [*refill*] de tinta (ISO 704, 2000, p. 8).

ANSI (2005)

Relação Todo-Parte: abrange situações em que um conceito está intrinsecamente incluso no outro, independentemente de contexto, de modo que os termos podem ser organizados em hierarquias lógicas, o todo pode ser tratado como um termo genérico (ANSI, 2005, p. 49).

Referentes:

Sistema nervoso; sistema nervoso central (cérebro; medula espinhal).

Khoo e Na (2006)

Relação Meronímia - Holonímia: trata a relação de parte-todo, mas enquanto na Hiponímia, a relação é interna aos conceitos, na Meronímia, a relação dá-se entre conceitos. [...] A relação meronímia é também apresentada como uma relação de parte-todo e partonímia e refere-se à relação entre um conceito/entidade e suas partes constituintes. A distinção entre as relações de meronímia e hiponímia é clara para os conceitos concretos, mas confusa para os conceitos abstratos. As relações de hiponímia podem vir a existir dentro de conceitos, enquanto as relações de meronímia estão entre conceitos [...] Relação parte-todo é intransitiva (KHOO, NA, 2006, p. 176).

Referentes:

Tipo de relação de parte-todo: 1. Componente funcional de um todo (por exemplo, roda de uma bicicleta); 2. O todo segmentado (o todo dividido em pedaços como uma torta); 3. Membros de uma coleção de elementos; 4. Subconjuntos de conjuntos (inclusão de grupo, por exemplo, frutas e maçãs) (IRIS, LITOWITZ; EVENS,

1988 apud KHOO; NA, 2007, p. 176).

Warner (1996 apud KHOO; NA, 2007, p. 177) listou as seguintes relações meronímicas: 1. membro-coleção (cachorro; matilha); 2. social whole/staff (avião; tripulação); 3. organização e seu chefe (navio; capitão); 4. um todo e sua unidade uniforme (areia; grão); 5. um todo e seu centro ou cume (montanha; pico [da montanha]).

Síntese dos enunciados – elementos do conceito “relação partitiva”

1. É uma relação ontológica.
2. É uma relação sintagmática.
3. É uma relação temporal.
4. É uma relação intransitiva.
5. A relação partitiva existe entre um todo e suas partes.
6. Os conceitos A e B são distintos.
7. O conceito superordenado representa um inteiro, enquanto o conceito subordinado representa parte deste inteiro.
8. As partes, juntas, formam um inteiro.
9. O conceito superordenado é chamado o conceito compreensivo e o conceito subordinado é chamado conceito partitivo.
10. Conceitos subordinados em um mesmo nível, que compartilham a mesma dimensão são chamados conceitos coordenados.
11. As partes de um objeto em um mesmo domínio podem variar.
12. Pode ocorrer da seguinte forma: a relação é interna aos conceitos (Hiponímia) e a relação dá-se entre conceitos (Meronímia). Partonímia refere-se à relação entre um conceito/entidade e suas partes constituintes.
13. O conceito que indica a parte pode ser essencial ou não essencial.

14. Relações partitivas aplicam-se a conceitos que expressam objetos.
15. O conceito referente à parte pode indicar uma propriedade físicas e intrínsecas de um objeto.
16. Tipo de relação de parte-todo: Componente funcional de um todo; o todo segmentado; Membros de uma coleção de elementos; Subconjuntos de conjuntos; social *whole/staff*; organização e seu chefe; um todo e sua unidade uniforme; um todo e seu centro ou cume.

D - RELAÇÃO DE CAUSA E EFEITO

Dahlberg (1978b) Relações funcionais aplicam-se sobretudo a conceitos que expressam processos. Pode-se conhecer o caráter semântico de tais relações tendo por base as chamadas valências semânticas dos verbos, dando atenção aos verbos e respectivos complementos (DAHLBERG, 1978b, p. 105).

Referentes: Produção – produto – produtos – produtor – comprador;
Objeto medido – fins da medição – instrumento de medição – graus de medição.

Farradane (1980b) Relação ação (*action*): quando os dois conceitos podem ser completamente discriminados e a associação entre os dois é temporária. Esta relação pode ser usada para descrever uma operação ou o efeito (“afetação”) de uma coisa sobre outra. Ex.: água/-purificação; decomposição/-prevenção; roupas/-detergente (FARRADANE, 1980b, p. 271).

Referentes: água/-purificação; decomposição/-prevenção; roupas/-detergente.

Farradane (1980b) Relação dependência funcional (*functional dependence*): quando

dois conceitos podem ser completamente discriminados, cuja associação entre eles é fixa, permanente. Expressa a relação de uma coisa causando ou produzindo algo ou mesmo um produto produzido por outra ‘coisa’. Aplica-se, especialmente, para a indexação de reações químicas (FARRADANE, 1980b, p. 272).

Referentes: Ex.: trigo /: pão; autor /: livro.

Khoo e Na (2006) Relação sintagmática. O conceito de causação é complexo e surpreendentemente difícil de definir. Podem se distinguir entre causas necessárias e causas suficientes. Um evento A é suficiente ainda que não uma condição necessária para o evento B se, quando A ocorre, B sempre segue, mas quando A não ocorre, B às vezes ocorre e às vezes não. A é uma necessidade ainda que não seja uma condição suficiente para B se, quando A não ocorre, B nunca ocorre, mas quando A ocorre, B às vezes ocorre ou às vezes não (KHOO, NA, 2006, p. 179).

Referentes: Não apresenta.

Síntese dos enunciados – elementos do conceito “relação de causa e efeito”

1. É uma relação ontológica.
2. É uma relação sintagmática.
3. É uma relação temporal.
4. É uma relação que indica processos.
5. É uma relação que indica ação.

E - RELAÇÕES ASSOCIATIVAS

Motta (1987) Uma relação que não se conformou com qualquer um dos anteriormente referidos, foi também identificada. Na falta de um nome mais adequado, tal relação foi chamada de ‘associação implícita’ (MOTTA, 1987, p. 50).

Referentes: Agricultura
TGR Setor primário (superordenação partitiva)
TR Alimentos (associação implícita) (MOTTA, 1987, p. 51).

Farradane (1980b) Relação associação (*association*): Quando dois conceitos são

difíceis de discriminar, cuja associação entre eles é fixa, permanente. Esta relação expressa várias formas de associação, as quais podem ser indefinidas (prisão /; desgraça), ou a relação com um agente de, ou ferramenta para, uma operação, etc. Expressa propriedades abstratas (obra de arte/; beleza) ou relações de ações no passado ou não especificadas (FARRADANE, 1980b, p. 271).

Referentes:

Prisão /; desgraça (indefinidas); obra de arte/; beleza (propriedades abstratas).

ISO 704 (2000)

Já as relações associativas são relações não hierárquicas. Existe uma relação associativa quando uma conexão temática pode ser estabelecida entre conceitos em virtude da experiência. Algumas relações associativas ocorrem quando a dependência é estabelecida entre conceitos com respeito a sua proximidade no espaço ou no tempo. Estas relações envolvem: matéria prima/produto; quantidade-unidade; material-estado; item concreto/material; item concreto/forma (*shape*); etc. Algumas relações envolvem eventos no tempo, como a dependência de um processo no tempo ou seqüência; outras relações envolvem causa e efeito (ISO 704, 2000, p. 10).

Referentes:

Matéria prima/produto;
Quantidade-unidade;
Material-estado;
Item concreto/material;
Item concreto/forma (*shape*);
Recipiente/o que é contido (estojo, lápis)
Atividade/ferramenta (escrever; lápis)
Passos de um ciclo (gametas; zigoto; zigosporo)
Produtor/produto (padeiro; pão);
Duração/equipamento de medida (tempo; relógio)
Profissão/ferramenta típica (pintor; pincel)
Objeto/ferramenta associada (parafuso; chave de fenda)
Organização/edificação associada (muçulmano; mesquita), etc.
(ISO 704, 2000, p. 11).

ANSI (2005)

Segundo a norma ANSI (2005, p. 51), a relação associativa é a mais difícil de definir e salienta que é importante explicitar a natureza das relações entre os termos e evitar julgamentos subjetivos, sob a pena de determinar relações associativas inconsistentes. Podem ocorrer entre termos de uma mesma hierarquia ou entre hierarquias diferentes. [...] Deve ser analisada, visto que não raro é considerada relação associativa tudo o que não é possível enquadrar como relação hierárquica ou equivalente.

Segundo a norma, algumas relações associativas ocorrem quando a

dependência é estabelecida entre conceitos com respeito a sua proximidade no espaço ou no tempo (ANSI, 2005, p. 12).

Uma relação entre dois ou mais termos em um vocabulário controlado que conduz de um termo a outros termos que são relacionados ou associados com ele (começa com as palavras VEJA TAMBÉM ou termos relacionados – TR) (ANSI, 2005, p. 157).

Referentes:

Processo/agente
 Processo/contra-agente
 Conceito ou objeto/propriedade
 Conceito ou objeto/origem
 Conceito ou objeto/ unidade ou mecanismo de medida
 Matéria bruta/produto
 Disciplina ou campo/objeto ou profissional (ANSI, 2005, p. 42).

Svenonius (2000)

Relações de termo relacionado: diretrizes tradicionais para a construção de linguagens de indexação definem as relações de termo relacionado vagamente, negativamente e amplamente para incluir todas as relações semânticas, exceto a equivalência e a hierarquia. A única propriedade matemática possuída pela relação de termo relacionado é a de simetria: se A está relacionada a B, então B está relacionada a A (SVENONIUS, 2000, p.160).

Referentes:

Não apresenta.

Síntese dos enunciados – elementos do conceito “relações associativas”

1. É uma relação ontológica.
2. É uma relação sintagmática.
3. É uma relação temporal.
4. É uma relação simétrica.

Nos quadros (QUADROS 10 a 14) a seguir são apresentados o mapeamento dos enunciados, por autor, das relações que originaram os enunciados descritos anteriormente.

a) Relações de equivalência - Neste grupo foram reunidos os tipos principais de equivalências (Sinonímia total; Variações lexicais; Quase-sinonímia ou Sinonímia-parcial) que consideram

conceitos muito próximos em sentido, inclusive a Antonímia, já que o entendimento de termos opostos compreende características comuns – mesmo que contrárias, ao invés de tratar termos opostos como conceitos associados, apenas (QUADRO 10).

b) Relações hierárquicas – Foram reunidas, de modo geral, as relações do tipo Gênero-espécie e de Instanciação, em que necessariamente o conceito genérico possui uma característica a menos que o termo específico, característica a qual, determina se tratar de uma espécie ou se de uma instância (ou uma instância da espécie). A identificação do conceito específico passa necessariamente pela identificação do conceito genérico, o que não ocorre em relações do tipo partitivas. Na Troponímia, tanto o termo genérico quanto o termo específico são verbos (QUADRO 11).

c) Relações partitivas - As relações partitivas foram consideradas à parte por se tratar de termos verificáveis apenas na realidade, portanto de ligação menos estável que nas relações hierárquicas. Segundo a literatura foram identificadas as seguintes categorias possíveis em relações deste tipo: Componente funcional/inteiro; Segmentos do inteiro; Membros/coleção; Subconjuntos de conjuntos; *Social whole/staff*; Organização e seu chefe; Inteiro e sua unidade uniforme; Inteiro/seu centro. A partir dos enunciados, identificamos as categorias: Ação/propriedade²¹; Conceito ou objeto/propriedade; material/propriedade e Matéria ou substância/propriedade. Estas estavam originalmente em relações associativas, mas de acordo com Farradane (1980b, p. 271), a relação de pertença indica inclusive, toda propriedade física intrínseca de um material ou coisa particular (QUADRO 12).

d) Relações de causa-efeito - segundo a literatura, principalmente por Farradane (ao indicar a *relação de dependência funcional*, quando dois conceitos podem ser completamente discriminados e expressam a relação de uma coisa causando ou produzindo algo ou mesmo um produto produzido por outra ‘coisa’; ou a *relação ação*, quando os dois conceitos podem ser completamente discriminados e a relação pode ser usada para descrever uma operação ou o efeito), foram identificadas as seguintes categorias possíveis para relações causa-efeito (originalmente empregadas em relações associativas): Ação/produto; Ação/propriedade; Ação/alvo; Ação/equipamento-ferramenta; Ação/lugar-localização e Ação/ator (QUADRO 12).

²¹ A categoria Ação/propriedade aparece como relação partitiva (ao se considerar propriedade como parte de uma ação) e como relação de causa-efeito (ao se considerar a propriedade desencadeada por uma ação).

e) Relações associativas – esta relação agruparia as categorias: Processo/agente; Processo/contragente; Conceito ou objeto/origem; conceito ou objeto/unidade ou mecanismo de medida; matéria bruta/produto; disciplina ou campo/objeto ou profissional; matéria prima/produto; Quantidade/unidade; Material/estado; Item concreto/material; Item concreto/forma; recipiente/o que é contido; Atividade/ferramenta; Passos de um ciclo; Produtor/produto; Duração/equipamento de medida; Profissão/ferramenta típica; Objeto/ferramenta associada; Organização/edificação associada (QUADRO 14).

Quadro 09 - Relação Semântica de Equivalência

RELAÇÃO SEMÂNTICA		AUTORES					
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesauros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
A - EQUIVALÊNCIA							
A1	Sinonímia total	Relação de identidade	Relação de equivalência	USE/UP	..	Relação de equivalência	Sinonímia
A2	Variações lexicais	Variações lexicais	..
A3	Quase-sinonímia ou Sinonímia-parcial	Relação de intersecção/ Inclusão	Quase-sinônimos	Quase-sinônimos/ Sinonímia-parcial
A4	Antonímia	Relação de oposição/ Disjunção	Relação de distinção	TO/Relação de oposição	..	Antonímia	Antonímia

Quadro 10 - Relação Semântica Hierárquica

RELAÇÃO SEMÂNTICA		AUTORES					
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesauros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
B - HIERÁRQUICA							
B1	Gênero-espécie	Relação hierárquica/ Inclusão	Relação de pertença	TG/TE	Relação genérica	Relação genérica	Hiponímia-hiperonímia
B2	Instanciação	Conceitos gerais e conceitos individuais	Relação de instanciação	..
B3	Troponímia	Troponímia

Quadro 12 - Relação Semântica Partitiva

RELAÇÃO SEMÂNTICA		AUTORES					
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) Tesauros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
C - PARTITIVA		Relação partitiva	Relação de pertença	TGP/TEP	Relação partitiva	Relação partitiva	Meronímia-holonímia/ Partonímia
	C1	Componente funcional / inteiro	Componente funcional - inteiro
	C2	Segmentos do inteiro	Segmentos do inteiro
	C3	Membros / coleção	Membros/ coleção
	C4	Subconjuntos de conjuntos	Subconjuntos de conjuntos
	C5	<i>social whole/staff</i>	<i>social whole-staff</i>
	C6	Organização/seu chefe	Organização e seu chefe
	C7	Inteiro e sua unidade uniforme	Inteiro e sua unidade uniforme
							(Continua...)

RELAÇÃO SEMÂNTICA		AUTORES					(Conclusão)	
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesauros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)	
(PARTITIVA)	C8	Inteiro/seu centro	inteiro/seu centro
	C9	Ação/propriedade	..	Relação de ação/ Relação de pertença	Relação atributiva	..	Ação/propriedade	..
	C10	Conceito ou objeto/propriedade	..	Relação de pertença	Relação atributiva	..	Conceito ou objeto/propriedade	..
	C11	Material/propriedade	..	Relação de pertença	Relação atributiva	Material/propriedade
	C12	Matéria ou substância-propriedade	..	Relação de pertença	Relação atributiva	Matéria ou substância - propriedade

Quadro 11 - Relação Semântica Causa-Efeito

RELAÇÃO SEMÂNTICA		AUTORES						
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesouros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)	
D - CAUSA-EFEITO		..	Relação de dependência funcional/ de ação	Relação de causalidade/ de afetação	Causa-efeito	Causa-efeito	Causa-efeito	
	D1	Ação/produto	..	Relação de dependência funcional/ de ação	Ação/produto	..
	D2	Ação/propriedade	..	Relação de dependência funcional/ de ação	Relação atributiva	..	Ação/propriedade	..
	D3	Ação/alvo	..	SIM(relação de ação)	..	Ação/alvo	Ação/alvo	
	D4	Ação-equipamento/ferramenta	..	SIM (relação de ação)	SIM (relação instrumental)	Ação-equipamento/ferramenta
	D5	Ação/lugar-localização	..	SIM (relação de ação)		Ação/lugar-localização
	D6	Ação/ator	..	SIM (relação de ação)		Ação/ator		

Quadro 12 - Relação Semântica Associativa

RELAÇÃO SEMÂNTICA		AUTORES (Cont.)					
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesaurus	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
E - ASSOCIATIVA		..	Relação de associação	TR/ Relação de associação implícita	Relação associativa	Relação associativa	..
E1	Processo/agente	Processo - agente	..
E2	Processo/contra-agente	Processo/co ntra-agente	..
E3	Conceito ou objeto/origem	Conceito ou objeto/origem	..
E4	Conceito ou objeto/ unidade ou mecanismo de medida	Conceito ou objeto/ unidade ou mecanismo de medida	..
E5	Matéria bruta/produto	Matéria bruta/produto	..
E6	Disciplina ou campo/objeto ou profissional	Disciplina ou campo/objeto ou profissional	..
E7	Matéria prima/produto	Matéria prima/produto

(Continua...)

RELAÇÃO SEMÂNTICA		AUTORES (Cont.)					
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesouros	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
(ASSOCIATIVA)							
E8	Quantidade-unidade	Quantidade -unidade
E9	Material/estado	..	Relação dimensional e estado	..	Material- estado
E10	Item concreto/material	Item concreto/m aterial
E11	Item concreto/forma	..	Relação dimensional e estado	..	Item concreto/fo rma
E12	Recipiente/o que é contido	..	Relação de pertença	..	Recipiente/ o que é contido
E13	Atividade/ferramenta	Relação instrumental	Atividade- ferramenta
E14	Passos de um ciclo	Passos de um ciclo
							(Continua...)

RELAÇÃO SEMÂNTICA		AUTORES					(Conclusão)
		Dahlberg (1978b)	Farradane (1980b)	Motta (1987) /Tesaurus	ISO 704 (2000)	ANSI (2005)	Khoo e Na (2006)
(ASSOCIATIVA)							
E15	Produtor/produto	Produtor – produto
E16	Duração/equipamento de medida	Relação instrumental	Duração- equipament o de medida
E17	Profissão/ferramenta típica	Relação instrumental	Profissão/ ferramenta típica
E18	Objeto/ferramenta associada	Relação instrumental	Objeto/ferr amenta associada
E19	Organização/edificação associada	Organização /edificação associada

Outro resultado esperado pela pesquisa, além de enumerar os tipos de relações semânticas, seria o de contribuir com a sistematização dos métodos de análise e proposição de relações em ciência da informação. Ao evidenciar as formas como se estabelecem os conceitos e relações em instrumentos terminológicos foi possível fazer alguns apontamentos para a identificação da forma como a modelagem conceitual pode ser determinada. Longe de se tornar um referencial para a modelagem conceitual das relações em ontologias, na próxima seção (4.2), são apresentadas breves considerações sobre os métodos identificados a partir das metodologias para o desenvolvimento de ontologias e das possíveis abordagens em ciência da informação, sendo as últimas, válidas para a construção de sistemas para a organização do conhecimento tradicionais.

4.2 Considerações sobre os métodos para estabelecimento de relações semânticas em ontologias

Ao verificar o tratamento dado às relações semânticas em ontologias, de acordo com a literatura pesquisada, pode-se apresentar a forma como ocorre a determinação das mesmas a partir das técnicas para o levantamento de conceitos; dos métodos de raciocínio utilizados e das garantias que podem validar este estabelecimento. Neste aspecto, privilegiou-se o olhar da ciência da informação, enunciando-se como esta pode se fazer presente na pesquisa em ontologias.

A determinação de conceitos e relações semânticas em ontologias é realizada na etapa de conceitualização, quando há a determinação do domínio e quando se define o escopo da ontologia, de acordo com os comprometimentos ontológicos que a linguagem assume. A delimitação dos conceitos válidos para a representação do domínio deve estar intimamente relacionada com a realidade que esta pretende representar.

Em síntese, os métodos pesquisados para a seleção de conceitos e relações em documentos envolvem a extração automática e/ou a extração semi-automática de conceitos e de relações. Para este tipo de levantamento terminológico – se automático, possível a partir do processamento por computador (processamento da linguagem natural em documentos disponíveis em ambiente digital) as principais fontes são os glossários, dicionários (LIM; SONG; LEE, 2004); os tesouros

(QIN; PALING, 2001; SOERGEL, *et al.* 2004 e FREITAS, *et al.*, 2008) ou ontologias já desenvolvidos (ALMEIDA, 2006) ou documentos disponíveis na *web*.

Fator de qualidade para a análise dos conceitos representativos de um domínio está na consideração da necessidade de compreender métodos “manuais” para o desenvolvimento de ontologias (LIM; SONG; LEE, 2004). O método de Gruninger e Fox (1995) utiliza a determinação de ‘questões de competência’ que a ontologia deverá ser capaz de responder, para capturar a terminologia de acordo com o dado contexto. Entre estas, há indicações que o levantamento por meio de técnicas para a extração de relações semânticas auxiliada por computador e o uso destas ferramentas não dispensa a análise de especialistas para a determinação de conceitos e relações válidos.

A distinção necessária entre instâncias, conceitos e relações é orientada tomando por base alguns critérios como: do agrupamento dos termos, como conceitos (seus atributos, instâncias e *constant*es – especificação de constantes, que especificam informações relacionadas ao domínio que sempre tem o mesmo valor) e como verbos, que representam ações no domínio, em Almeida (2005, p. 17). Para Guarino e Welty (2002, p. 63) validarem a estrutura taxonômica de uma ontologia, os autores propõem a utilização de ‘noções formais’ de essência e rigidez; identidade e unidade.

Para a distinção de instâncias, por exemplo, a teoria do conceito apresenta os conceitos de conceitos/objetos individuais e conceitos/objetos gerais que se aplicam a análise dos conceitos e auxiliam nessa distinção.

Toda vez que o objeto é pensado como único, distinto dos demais, constituindo uma unidade inconfundível (coisas, fenômenos, processos, acontecimentos, atributos, etc.) pode-se falar de *objetos individuais*. Pode-se dizer que o que caracteriza os objetos individuais é a presença das formas do tempo e espaço [...]. Mas, além dos objetos individuais, expressos pelos conceitos individuais, podemos referir-nos a objetos gerais que, de certo modo, prescindem das formas do tempo e espaço. A esses objetos situados fora do tempo e do espaço, correspondem os chamados conceitos gerais (DAHLBERG, 1978b, p. 101-102).

Em métodos que se utilizam da transposição de tesouros já desenvolvidos para a construção de ontologias (QIN; PALING, 2001; SOERGEL, *et al.* 2004 e FREITAS, *et al.*, 2008), a escolha das

relações não pode ocorrer pela simples transposição de um suporte ao outro, mas prescinde de uma adequação ao domínio da ontologia, já que nos primeiros, as relações se limitam a três tipos.

Nas metodologias para ontologias, é possível identificar que a determinação dos conceitos e relações ocorre por processo de generalização e de especialização (USCHOLD; KING, 1995; GANDON, 2001; HOVY, 2003). O processo de generalização ou *bottom-up* inicia com a definição dos conceitos mais específicos de um domínio, seguido do agrupamento dos mesmos em classes. O processo de especialização ou *top-down* inicia com a definição dos conceitos mais gerais de um domínio, seguida da especialização dos mesmos. Fora o privilégio na utilização de um destes métodos de raciocínio, o que ocorre é a combinação dos dois métodos – o *middle-level*, quando os processos de generalizar e especializar ocorrem simultaneamente.

Nenhum destes três métodos é intrinsecamente melhor do que qualquer dos outros. A abordagem a adotar depende fortemente do ponto de vista pessoal do domínio. Se um desenvolvedor tem uma visão ampla e sistemática do domínio, então pode ser fácil usar a abordagem *top-down*. A combinação das abordagens é com frequência a mais fácil para muitos desenvolvedores, uma vez que os conceitos “no meio” tendem a serem conceitos mais descritivos no domínio²² (NOY, GUINNESS, 2001, p. 7).

Segundo Campos (2004, p. 25), método *top-down* é também denominado método dedutivo e o método *bottom-up* é também denominado método indutivo em ciência da computação. Em ciência da informação, os métodos dedutivo e indutivo estão presentes para a identificação de conceitos de um dado domínio, para fins de construção de tesouros e sistemas bibliográficos.

O método dedutivo, para a elaboração de sistemas de classificação, consiste na subdivisão de um universo do conhecimento em disciplinas (MOTTA, 1987, p. 25). Nessa abordagem, o conjunto de termos é definido *aprioristicamente*, como uma rede de relações entre significados dos termos, sendo essa rede construída via análise conceitual. Dodebei (2002, p. 70) acrescenta que no método de dedução, a obtenção da terminologia se faz por consenso de especialistas peritos no assunto.

²² No original, “None of these three methods is inherently better than any of the others. The approach to take depends strongly on the personal view of the domain. If a developer has a systematic top-down view of the domain, then it may be easier to use the top-down approach. The combination approach is often the easiest for many ontology developers, since the concepts “in the middle” tend to be the more descriptive concepts in the domain.” (NOY, GUINNESS, 2001p. 7).

Segundo Dodebei (2002), no processo indutivo, faz-se a obtenção da terminologia mediante a identificação de termos prováveis, a partir da frequência de ocorrência dos termos no exame da literatura corrente: [...] onde as árvores hierárquicas são construídas a partir dos termos coletados, até onde se obter máximos gêneros e mínimas espécies, que darão origem às classes gerais ou facetas principais (DODEBEI, 2002, p. 70).

Motta (1987, p. 27) já indicava esse princípio (garantia literária e garantia do usuário), como comuns à construção de sistemas de classificação em geral, mas ressalta que esses são referidos na literatura mais comumente para a seleção dos termos do que para a identificação das relações – ponto contemplado em Barité (2007) e Hjørland (2007c).

As garantias que são retomadas brevemente nesta seção são originalmente tratadas em Begthol (1986). Podem ser de várias naturezas²³, mas optou-se por tratar apenas as seguintes - por melhor se adequarem ao verificado em metodologias para a construção de ontologias: a garantia literária (o papel da análise documental); a garantia do usuário (para validação das definições dos conceitos junto aos especialistas e usuários finais) e a garantia cultural (a importância dos comprometerontes ontológicos para a representação do domínio).

Barité (2007) apresenta a garantia literária como uma ferramenta útil para distinguir entre relações paradigmáticas e sintagmáticas “e para legitimar aquelas relações paradigmáticas que por sua força e estabilidade deveriam estabelecer-se em um tesouro ou outra linguagem documentária”. Considera, então, a validade do levantamento de relações semânticas além dos conceitos recorrentes na literatura (de um domínio). Para ele, a garantia literária está consolidada no domínio da organização do conhecimento, como uma formulação teórica e metodológica necessária para a construção e manutenção de estruturas conceituais de qualquer espécie e finalidade, que buscam representar a dinâmica (de conceitos e de relações paradigmáticas) de um domínio do conhecimento²⁴.

²³ Garantia filosófica; garantia educacional, para citar algumas.

²⁴ No original, “[...]y el mantenimiento de estructuras conceptuales de cualquier especie y finalidad, que busquen estar de acuerdo con el devenir de los dominios del saber metodología de legitimación de la terminología seleccionada para servir de nexo entre la información y los usuarios, y se propone comprender en la noción de garantía literaria no sólo la justificación de términos, sino también la justificación de las relaciones paradigmáticas que se crean pertinentes en un determinado sistema de información” (BARITÉ, 2007, [p.8]).

Ainda para o desenvolvimento e a validação do modelo em ontologias, além da análise de documentos devem-se considerar informações para o domínio com atores/usuários do domínio. Para o desenvolvimento de ontologias, podemos citar as seguintes técnicas e ferramentas para capturar essa validação: entrevistas com especialistas; observação participante; *brainstorming*, etc, que, apesar de não detalhadas na presente pesquisa, se mostram relevantes para este tipo de levantamento. Em ciência da informação, pode-se fazer um paralelo com a “Garantia do usuário”, que corresponde à validação das definições dos conceitos junto aos especialistas (já a garantia de uso, trata da efetividade de uso dos conceitos/relações para os usuários finais²⁵). Esta garantia, junto à garantia cultural (ao considerar a representação de diferentes opiniões epistemológicas e culturais, por exemplo, de minorias), se aplica perfeitamente para o desenvolvimento de ontologias.

Metodologicamente, estas garantias podem validar, o respeito ou não, aos comprometimentos ontológicos assumidos para a representação de uma dada realidade. Cujas limitações é relembrada em Campos (2004), ao considerar que

uma vez que toda representação é uma aproximação imperfeita da realidade, ao selecionarmos uma representação, estamos tomando um conjunto de decisões sobre como e o que ver no mundo. Portanto, selecionar uma representação significa fazer um conjunto de compromissos ontológicos. Esses compromissos determinam o que pode ser visto, enfocando alguma parte do mundo em detrimento de outras (CAMPOS, 2004, p. 24, grifo nosso).

A partir da análise do tratamento dado às relações semânticas em ontologias, pode-se afirmar que as técnicas, métodos de raciocínio e garantias apresentados não se sobrepõem, mas podem e devem ser intercalados na metodologia, se o objetivo da metodologia é efetivamente o de desenvolver ferramentas expressivas da realidade que se pretende modelar. A soma destas ações pode resultar em análises mais robustas na determinação de conceitos e de relações semânticas apropriados em ontologias.

²⁵ Em tesouros, por exemplo, a indicação dos termos preferidos e os não preferidos deve levar em consideração a frequência de solicitação do conceito pelos usuários.

5 CONSIDERAÇÕES FINAIS

A presente pesquisa partiu do estudo das relações semânticas para ampliar a compreensão da natureza das mesmas e para colaborar metodologicamente (na fase de modelagem conceitual) para o desenvolvimento de ontologias, cuja temática é de interesse para a ciência da informação. Tradicionalmente, esta área de conhecimento se ocupa da natureza dos conceitos e de relações entre os conceitos para a construção de ferramentas para a organização do conhecimento, como verificado na literatura. A partir do desenvolvimento e resultados deste trabalho, retornar-se-á às questões e as hipóteses que encadearam a pesquisa.

Inicialmente, as questões que instigaram a pesquisa, indagavam sobre a possibilidade de enumerar e identificar as relações semânticas, *a priori*. Considerando-se que as relações semânticas têm forte ligação com o próprio crescimento do conhecimento, podemos inferir a impossibilidade de enumerá-las *a priori*, ou de pelos menos esgotá-las. Mesmo com esta limitação, neste trabalho, chegou-se à seguinte sistematização das relações semânticas em ferramentas para a OC, apresentadas em 5 grupos: relações de equivalência; relações hierárquicas; relações partitivas; relações de causa e efeito e relações associativas.

O trabalho, no entanto, não pretendeu esgotar todos os tipos de relações possíveis por compreender que não é possível determinar *a priori*, todas as relações semânticas possíveis. Inclusive por considerar que, mesmo o levantamento aqui apresentado não é definitivo – outros trabalhos podem ser incorporados. A impossibilidade de esgotar todas as possibilidades de relacionamento não significa, portanto, que o levantamento das relações existentes em um determinado domínio seja inválido ou um esforço inútil. Este impedimento, ao contrário, reforça a necessidade da análise conceitual para a identificação das relações semânticas em domínios e aplicações específicos.

Afinal, conclui-se que mais importante que propor uma estrutura de relações para a construção de ontologias, o que podemos apontar como o resultado mais expressivo da presente pesquisa é a aplicação do método (de análise conceitual) para a “desconstrução” das relações de modo a tratar sua dinâmica em um dado domínio. A existência de um núcleo de relações comuns a todas as áreas do conhecimento não substituiria a análise da estrutura conceitual do domínio em si.

As hipóteses levantadas inicialmente neste trabalho estão retomadas nestas considerações finais. A primeira hipótese considerada nesta pesquisa, seria a de que o aporte teórico e metodológico em ciência da informação seria suficientemente robusto para a pesquisa e desenvolvimento de ontologias, justamente por envolver outras áreas do conhecimento. Das contribuições desta área de conhecimento, podemos citar as que diretamente identificamos:

- a) O uso combinado de técnicas, métodos de raciocínio e garantias amplamente utilizadas em ciência da informação, para o desenvolvimento e validação de instrumentos terminológicos, podem ser aplicados para a construção de ontologias.
- b) Propostas anteriores de sistematização das relações semânticas (distinção paradigmática e sintagmática; relações *a priori* e *a posteriori*; relações lógicas e ontológicas, relações permanentes (atemporais) e relações temporárias, a pré-coordenação dos conceitos e a indexação relacional, propriamente,) apontam que a área tem desenvolvido corpo teórico que se ocupa da natureza e da dinâmica das relações para a representação do conhecimento.

A segunda hipótese considera que indiferente à mudança de ambiente – dos acervos físicos para os acervos digitais – mudam-se as ferramentas para se realizar os processos de armazenamento, tratamento e recuperação da informação, mas o esforço intelectual empreendido, os processos dos quais se ocupa a organização do conhecimento e as motivações para a organização não sofreram alterações estruturais. O que ocorreu foi o aprimoramento das próprias ferramentas - que coincide com a terceira hipótese da pesquisa, ao compreender que, com o desenvolvimento da pesquisa em ontologias, tem-se instrumental necessário para a aplicação de estudos anteriores, visando-se a efetiva aplicação dos estudos para o aprimoramento das técnicas para a organização do conhecimento.

Para trabalhos futuros podemos, afinal, indicar:

- a) Verificação da aplicação das relações semânticas extraídas da literatura da ciência da informação para o desenvolvimento de ontologias. Com isso, têm-se condições

de verificar a relevância destas relações e a conveniência dos elementos dos conceitos apontados nos enunciados da presente pesquisa.

- b) Aprofundar a pesquisa em semântica e linguística, de forma a explicitar sua inserção em ciência da informação, em especial para o desenvolvimento de ontologias.
- c) Verificação de ambigüidades semânticas em ontologias ocasionadas por relações semânticas indevidas.
- d) Um desdobramento necessário desta pesquisa é, a partir do esclarecimento conceitual das relações semânticas, subsidiar pesquisas posteriores para a formalização das mesmas. Para isso é necessário o estudo de linguagens específicas para o desenvolvimento de ontologias, de modo a garantir menor perda semântica possível da relação em linguagem natural ao seu equivalente em linguagem formal.

REFERÊNCIAS

ABNT - Associação Brasileira de Normas Técnicas. **NBR 12676**: Métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

AGROVOC. Disponível em: http://www.fao.org/aims/ag_intro.htm. Acesso em: 28 jul. 2009.

ALMEIDA, M. B. **Metodologia para construção de modelos corporativos baseados em ontologias**: minicurso ministrado no II Simposio Mineiro de Sistemas de Informação – SMSI. Belo Horizonte: Faculdade COTEMIG, 2005. 46p.

_____. **Um modelo baseado em ontologias para representação da memória organizacional**. 2006. 316f. Tese (Doutorado em ciência da informação) – Escola de ciência da informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2006. Disponível em: <<http://hdl.handle.net/1843/VALA-6T7QFT>>. Acesso em: 28 jul. 2009.

ALVES, José Wagner Vasconcelos. **Evolui**: uma metodologia para auxiliar a criação de ontologias. 2003. 126f. Dissertação (Mestrado em informática aplicada) – Universidade de Fortaleza, Fortaleza, 2003.

AMERICAN NATIONAL STANDARDS ORGANIZATION. ANSI/NISO Z 39.19: 2005: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005. 176 p. Disponível em: <<http://www.niso.org/standards/resources/Z39-19-005.pdf?CFID=2436046&CFTOKEN=81298864>>. Acesso em: 27 jun. 2007.

ANTONIOU, G. **OWL**: web ontology language. In: ANTONIOU, G.; VAN HARMELEN, F. (Ed.). *A semantic web primer*. Cambridge: MIT, 2004. p. 109-150.

AUSTIN, D. The development of PRECIS: a theoretical and technical history. **Journal of documentation**, v. 30, n. 1, p. 97-102, 1974.

BARITÉ, M. La garantía literaria: vigencia y proyección teórico-metodológica. In: Encontro Nacional de Pesquisadores em ciência da informação e Biblioteconomia. 8. **Anais...** 2007. Salvador: UFBA, 2007.

BEDIN, Sonali Paula Molin. **Metodologia para validação de ontologias: o caso ORBIS_MC**. 2007. 113f. Dissertação (Mestrado em ciência da informação) – Programa de Pós-Graduação em ciência da informação, Universidade Federal de Santa Catarina, Florianópolis, 2007.

BEGHTOL, Clare. Semantic validity: concepts of warrant in bibliographic classification systems. **Library resources and technical services**, v. 30, n. 2, p. 109-25, 1986.

BRASCHER, M; CAFÉ, L. Organização da informação ou organização do conhecimento? In: Encontro Nacional de Pesquisadores em ciência da informação e Biblioteconomia. 9. **Anais...** 2008, São Paulo: USP. 2008.

BREITMAN, Karin. **web semântica: a internet do futuro**. Rio de Janeiro: LTC, 2005.

BROOKES, B. C. Jason Farradane and relational indexing. **Journal of Information Science**. v. 12, n. 1-2, p. 15-18, 1986.

CAMPOS, Maria Luiza de Almeida. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: UFF, 2001.

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

_____; 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. **Ciência Informação**, Brasília, v. 32, n. 2, maio/ago. 2003.

CERQUEIRA, Roberto Figueiredo Paletta de. **Método de modelagem domínio-ontológica do Direito Positivo Brasileiro**. 2007. 133f. Dissertação (em ciência da informação) – Escola de ciência da informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2007. Disponível em: <<http://hdl.handle.net/1843/VALA-74QKF5>>. Acesso em 28 jul. 2009.

CLARKE, S.G. D. Thesaural relationships. In C.A. BEAN; R. GREEN (Eds.). **Relationships in the organization of knowledge**. Dordrecht: Kluwer, 2001. p. 37-52.

CLASSIFICATION RESEARCH GROUP. The need for a faceted classification as the basis of all methods of information retrieval. UNESCO document 320/5515 (International advisory committee for documentation and terminology in pure and applied science). Paris, 1955 *apud* FROHMANN, B. P. An investigation of the semantic bases of some theoretical principles of

classification proposed by Austin and the CRG. **Cataloging & Classification Quarterly**, v. 4, n. 1, 11-27. 1983.

COATES, E. J. Some properties in the structure of indexing languages. **Journal of documentation**, London, v. 29, n. 4, p. 390-404, dez. 1973.

CROS, R. C; GARDIN, J. C; LEVY, F. **L'Automatisation des recherches documentaires**: un modèle général, le Syntol. 2. ed. Paris: Gauthier-Villars, 1968.

CRUSE, D. A. **Lexical semantics**. Cambridge: Cambridge University Press. 1986.

CUNHA, Isabel M. R. Análise documentária. In: SMITH, J. (Coord.). **Análise documentária**: análise da síntese. Brasília: IBICT, 1987. cap. 4, p. 38-60.

DAHLBERG, I. A referent-oriented analytical concept theory of interconcept, **Intern.Classificat.**, Frankfurt, v. 5, n. 3, p. 142-150, 1978a.

_____. **Conceptual structures and sistematization**. International Forum on Information and documentation, v. 20, n. 3, p.09-24, Jul. 1995

_____. Futuro das linguagens de indexação. In: Conferencia Brasileira de Classificação Bibliográfica, 1976, Rio de Janeiro. **Anais...** Comemorativa do centenário da primeira edição de Classificação Decimal de Dewey. Rio de Janeiro: IBICT; Associação dos Bibliotecarios do Distrito Federal. 1979. p. 323-334. Disponível em: <<http://www.conexaorio.com/bit/dahlberg/index.htm>>. Acesso em: 28 jul. 2009.

_____. Knowledge organization: its scope and possibilities. **Knowledge Organization**, v. 20, n. 4, 211-222, 1993.

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

DIAS, E. W. Contexto Digital e Tratamento da Informação. **DataGramZero - Revista de ciência da informação** - v. 2, n. 5, out. 2001. Disponível em: <http://dgz.org.br/out01/Art_01.htm>. Acesso em: 28 jul. 2009.

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

FARRADANE, J. E. L. A scientific theory of classification and indexing and its practical application. **The Journal of Documentation**. v. 6, n. 2, p. 83-99, 1950.

_____. A scientific theory of classification and indexing: further considerations. **The journal of documentation**, v. 8, n. 2, p. 83-99, jun. 1952.

_____. Concept organization for information retrieval. **Information Storage and Retrieval**, 3(4), 297-314, 1967.

_____. Knowledge, information, and information science. **Journal of information science**, v. 2, p. 75-80, 1980a.

_____. Relational Indexing: part I. **Journal of information science**, n.1, p. 267-276, 1980b.

_____. Relational indexing: part II. **Journal of information science**, n. 1, p. 313- 324, 1980c.

_____. The psychology of classification. **Journal of Documentation**, v. 11, n. 4 , p. 187- 201, 1955.

FERNANDEZ, M.; GOMEZ-PEREZ, A.; JURISTO, H. **Methontology; from ontological art towards ontological engineering**. 1997. Disponível em: < <http://www.aaai.org/Library/Symposia/Spring/1997/ss97-06-005.php>>. Acesso em: 28 jul. 2009.

FILLION, et al. **An ontology-based environment for enterprise model integration**. (1995). Disponível em: < <http://citeseer.ist.psu.edu/context/400800/0>>. Acesso em: 04 dez 2008.

FONSECA, F. The double role of ontologies in Information Science research. **JASIST**, v. 58, n. 6, p. 786-793, 2007.

FOSKETT, A. C. **A abordagem temática da informação**. São Paulo: Polígono, 1973. 437p.

FOX, M. S.; GRUNINGER, M. Enterprise modeling. **AI Magazine**, v. 19, n. 3, p.109-121, 1998. Disponível em: <<http://www.eil.utoronto.ca/enterprise-modelling/papers/fox-aimag98.pdf>>. Acesso em: 29 jul. 2009.

FRANÇA, Júnia Lessa. Manual para normalização de publicações técnico-científicas. 7. ed. Belo Horizonte: UFMG, 2004.

FREITAS, K. F. ; CAMPOS, M. L. M. ; CAMPOS, M. L. A. ; CAMPOS, L. M. . Tesouro como base terminológica para a elaboração de ontologia de domínio: uma aplicação no domínio do Folclore e Cultura Popular. In: Seminário de Pesquisa em Ontologia no Brasil, 2008, Niterói. **Anais...**, 2008. Disponível em: <<http://www.uff.br/ontologia/artigos/323.pdf>>. Acesso em: 29 jul. 2009.

FROHMANN, B. P. An investigation of the semantic bases of some theoretical principles of classification proposed by Austin and the CRG. **Cataloging & Classification Quarterly**, v. 4, n. 1, 11-27. 1983.

GANDON, F. Engineering an ontology for a multi-agents corporate memory system. In: PROC. INTERNATIONAL SYMPOSIUM ON THE MANAGEMENT OF INDUSTRIAL AND CORPORATE KNOWLEDGE, p.209-228, 2001. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.21.4772>>. Acesso em: 29 jul. 2009.

GARCIA JIMÉNEZ, A. Instrumentos de representation del conocimiento: tesauros versus ontologias. **Anales de documentacion**, n. 7, p. 79-95, 2004.

GARDIN, Jean-Claude. Document analysis and linguistic theory. **Journal of documentation**, London, v. 29, n. 2. p. 137-168, jun. 1973.

GHOSE, A. DHAWLE, A. S. Problems of thesaurus construction. **Journal of the American society for information science**, v. 28, n. 4, p. 211-217, 1977.

GILCHRIST, Alan. Thesauri, taxonomies and ontologies: an etymological note. **Journal of documentation**, v. 59, n. 1, 2003, p. 7-18. Disponível em: <<http://dois.mimas.ac.uk/DoIS/data/Articles/julkokltny:2003:v:59:i:1:p:7-18.html>>. Acesso em: 14 nov. 2006.

GOMES, H. E. Biblioteconomia, Informação e Tecnologia da Informação. **Elaboração de tesauros documentários**: tutorial. Disponível em: <<http://www.conexaorio.com/bit/tesauro/index.htm>>. Acesso em: 29 jul. 2009.

_____. (Org.). **Manual de elaboração de tesauros monolíngues**. Brasília: Programa Nacional de Bibliotecas de Instituições de Ensino Superior, 1990.

GREEN , P., ROSEMANN. Editorial preface: Ontological analysis, evaluation and engineering of business systems analysis methods. **Journal of Database Management**, v. 15, Apr./Jun, 2004. Disponível em: <[http://igi-pub.com/files/prefaces/jdm_preface_15\(2\).pdf](http://igi-pub.com/files/prefaces/jdm_preface_15(2).pdf)>. Acesso em: 29 jul. 2009

GRUBER, T. **A translation approach to portable ontology specifications**. Knowledge Acquisition, v. 5, p. 199-220, 1993. Disponível em: <http://www-ksl.stanford.edu/KSL_Abstracts/KSL-92-71.html>. Acesso em: 29 jul. 2009.

_____. **What is an Ontology?** 1996. Disponível em: <<http://www.ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: 29 jul. 2009.

GUARINO, N. Formal Ontology and Information Systems. In: GUARINO, N. (Ed.) **Formal Ontology in Information Systems**. p. 3-15, Amsterdam: IOS Press, 1998. Disponível em: <<http://www.loa-cnr.it/papers/fois98.pdf>> Acesso em: 29 jul. 2009.

_____. **Understanding, building, and using ontologies**. 1996. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/guarino.html>>. Acesso em: 29 jul. 2009.

_____; GANGEMI, Aldo. **New Classification of Ontologies**. [2005?]. Disponível em: <<http://www.aifb.uni-karlsruhe.de/WBS/dob/book/highlights.html>>. Acesso em: 29 jul. 2009.

_____; GIARETTA, P. Ontologies and Knowledge bases, towards a terminological clarification. In: MARS, N. (Ed.). **Towards a Very Large Knowledge Bases; Knowledge Building and Knowledge Sharing**. [S.l.]: IOS Press, 1995. p. 25-32. Disponível em: <<http://www.loa-cnr.it/Papers/KBKS95.pdf>>. Acesso em: 29 jul. 2009.

_____; WELTY, C. **Evaluating ontological decisions with OntoClean**. **Communications of the ACM**, v. 45, n. 2, fev. 2002. Disponível em: <<http://portal.acm.org/citation.cfm?id=503150>>. Acesso em: 30 jul. 2009.

HJORLAND, B. **Discussion of Dahlberg's theory of concepts and knowledge organization**. 2007a. Disponível em: <http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/discussion_of_dahlberg.htm>. Acesso em: 29 jul. 2009.

_____. **Semantic relations: meaning relations**. 2007b. Disponível em: <http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/semantic_relations.htm>. Acesso em: 29 jul. 2009.

_____. Semantics and Knowledge Organization. **Annual Review of Information Science and Technology**, v. 41, p. 367-405, 2007c.

HODGE, G. **Systems of knowlege organization for digital libraries**: beyond traditional authority files (CLIR Report 91). Washington, DC: Council on Library and Information Resources. Disponível em: <www.clir.org/pubs/reports/pub91/pub91.pdf>. Acesso em: 30 jul. 2009.

INTERNATIONAL STANDARD ORGANIZATION. **ISO 704**: Terminology work: Principles and methods principles and methods of terminology. Geneva: ISO, 2000.

IRIS, M. A; LITOWITZ, B. E; EVENS, M. Problems of the part-whole relation. In: EVENS, M. W. **Relational models of the lexicon**: Representing knowledge in semantic networks. Cambridge: Cambridge University Press, 1988, Cap. 12, p. 261-288.

JANSEN, R. Die Bedeutung der Zugehörigkeitsrelation für Thesaurusstruktur, Indexierungsmethode und Recherche. In: **Deutscher Dokumentartag**, 1974. Band 1, Bonn, 7-11 Okt 1974 / Deutsche Gesellschaft für Dokumentation. Munkun: Verlag Dokumentation, 1975. p. 353-374 *apud* RIVIER, Alexis. Construção de linguagens de indexação: aspectos teóricos. **Revistada Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 21, n. 1, p. 56-99, jan./jun. 1992.

JAPIASSÚ, H.; MARCONDES, D. **Dicionário básico de filosofia**. Rio de Janeiro: Jorge Zahar, 2006. 300p.

JASPER, R.; USCHOLD, M. **A Framework for understanding and classifying ontology applications**. 1999. Disponível em: <<http://sern.ucalgary.ca/KSI/KAW/KAW99/papers/Uschold2/final-ont-apnfmk.pdf>>. Acesso em: 12 jul. 2005.

KHOO, C. S. G.; NA, Jin-Cheon. **Semantic relations in Information Science**. Annual Review of Information Science and Technology, v. 40, p. 157-228, 2006.

LANCASTER, F. W. **Vocabulary control for information retrieval**. Arlington, VA: Information Resources Press, 1986.

_____. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 2004.

LE COADIC, Y.F. **A ciência da informação**. Brasília, DF, Briquet de Lemos. 2004.

LEGG, C. Ontologies on the Semantic Web. **Annual Review of Information Science and Technology**. v. 41, p. 407-491, 2007.

LYONS, J. **Linguistic semantics**: an introduction. Cambridge: Cambridge University Press. 1995.

MANIEZ, Jacques. Relationships in thesauri: some critical remarks. **International classification**, v. 15, n. 3, 1988. p. 133-138.

MOREIRA, A. Contribuição da terminologia na modelagem de sistemas computacionais. **DataGramaZero - Revista de ciência da informação** - v. 6, n. 5 out./05. Disponível em: <http://dgz.org.br/out05/Art_01.htm>. Acesso em: 30 jul. 2009.

_____. **Tesauros 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. 150 p. Dissertação (Mestrado em ciência da informação)-Escola de ciência da informação da UFMG, Universidade Federal de Minas Gerais, Belo Horizonte, 2003.

MOTTA, Dilza Fonseca da. **Método relacional como nova abordagem para a construção de tesauros**. 1987. 89 f. Dissertação (Mestrado em ciência da informação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1987.

MYAENG, S. H., JANG, D. Development and evaluation of a statistically based document summarization system. In I. MANI; M. T. MAYBURY, M. T., **Advances in automatic text summarization**. Cambridge, MA: MIT Press, 1999. p. 61-70.

_____; MCHALE, M. L. Toward a relation hierarchy for information retrieval. In B. H. KWASNIK; R. FIDEL, **Advances in Classification Research: Proceedings of the 2nd ASIS SIG/CR Classification Research Workshop**. Medford, NJ: Learned Information. 1992, p. 101-113.

NOY, F. N.; GUINNESS, D. L. **Ontology development 101**: a guide to create your first ontology. 2001. Disponível em: <http://protege.stanford.edu/publications/ontology_development/ontology101.pdf>. Acesso em: 30 jul. 2009.

OBO Relation Ontology. Disponível em: <www.obofoundry.org/ro/>. Acesso em: 14 jul. 2007.

OGDEN, C. K.; RICHARDS, I. A. **O significado de significado**: um estudo da influência da linguagem sobre o pensamento e sobre a ciência do simbolismo. Rio de Janeiro: Zahar, 1972. 349 p.

OLIVEIRA, Káthia M. **Modelo para construção de ambientes de desenvolvimento de software orientados a domínio**. Tese de doutorado, UFRJ, Rio de Janeiro, 1999 *apud* ALVES, José Wagner Vasconcelos. **Evolui**: uma metodologia para auxiliar a criação de ontologias. 2003. 126f. Dissertação (Mestrado em informática aplicada) – Universidade de Fortaleza, Fortaleza, 2003.

OLIVEIRA, Roberta Pires de. Semântica. In: MUSSALIM, Fernanda; BENTES, Anna Christina. **Introdução à lingüística**: domínios e fronteiras. 2.ed. v.2. 2000. p. 17-46.

ONTOLOGIA. In: DICIONÁRIO Houaiss: dicionário eletrônico da língua portuguesa. Rio de Janeiro, Ed. Objetiva, 2001.

OWL: web ontology language reference. Disponível em: <<http://www.w3.org/TR/owl-ref/>> Acesso em: 30 jul. 2009.

PERAKATH, B. C. et al. **IDEF5 Method Report**. 1994. Disponível em: <<http://www.idef.com/pdf/Idef5.pdf>>. Acesso em: 30 jul. 2009.

PIEDADE, M. A. R. **Introdução à teoria da classificação**. 2. ed. Rio de Janeiro: Interciência, 1983. 221 p.

QIN, J.; PALING, S. Converting a controlled vocabulary into an ontology: the case of GEM. 2001. **Information Research**, v. 6, n. 2. Disponível em: <http://InformationR.net/ir/6-2/paper94.html>. Acesso em: 23 jun. 2005.

RANGANATHAN, S. R. **Prolegomena to library classification**, Bombay: Ásia Publishing House, 1967. 640p *apud* CAMPOS, Maria Luiza de Almeida. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: Ed. da UFF, 2001.

_____. **The Colon Classification**. New Brunswick, N.J.: Graduate School of Library Service, Rutgers University. 1965.

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

SALES, L. F. **Modelo triádico de relações para aplicação em Ontologias** - seminário de pesquisa em ontologia no brasil. 2007a.

_____. **Relações conceituais para instrumentos de padronização terminológica: um novo modelo para o uso em ontologias**. 2007b. Disponível em: <<http://www.enancib.ppgci.ufba.br/artigos/GT2--088.pdf>>. Acesso em 30 jul. 2009.

_____. **Ontologias de domínio**: estudo das relações conceituais e sua aplicação. 2006. 141 f. Dissertação (Mestrado em ciência da informação) – Universidade Federal Fluminense, Niterói, 2006.

SARACEVIC, T. ciência da informação: origem, evolução, relações. **Perspectivas em ciência da informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SELLTIZ, C. et al. **Métodos de pesquisa nas relações sociais**. São Paulo: Herder, 1967.

SILVA, D.L. da; SOUZA, R.R.; ALMEIDA, M.B. Ontologias e vocabulários controlados: comparação de metodologias para construção. **Ciência da informação**, Brasília, v. 37, n. 3, p. 60-75, set./dez. 2008.

SMITH, B. **Ontology**. 2003. Disponível em: <http://ontology.buffalo.edu/smith/articles/ontology_pic.pdf>. Acesso em: 30 jul. 2009.

_____. *et al.* **Relations in biomedical ontologies**. *Genome Biology* 6. 2005. Disponível em: <<http://genomebiology.com/2005/6/5/R46>>. Acesso em: 30 jul. 2009.

SOERGEL, D. *et al.* Reengineering Thesauri for New Applications: the AGROVOC example. **Journal of Digital Information**, 2004. Disponível em: <<http://www.fao.org/docrep/008/af234e/af234e00.htm>>. Acesso em: 30 jul. 2009.

SOUZA, S. de. **CDU**: como entender e utilizar a edição-padrão internacional em língua portuguesa. 3. ed. rev. atual. Brasília: Thesaurus, 2004.

SOWA, J. F. **Knowledge representation**: logical, philosophical, and computational foundations. Pacific Grove: Brooks-Cole, 2000. 594 p.

SUBSUNÇÃO. In: DICIONÁRIO Houaiss: dicionário eletrônico da língua portuguesa. Rio de Janeiro, Ed. Objetiva, 2001.

SVENONIUS, E. The intellectual foundation of information organization. Cambridge, MA: MIT Press, 2000.

SWARTOUT, B. et al. Toward Distributed Use of Large-Scale Ontologies. In: PROC. OF AAAI97 SPRING SYMPOSIUM SERIES WORKSHOP ON ONTOLOGICAL ENGINEERING. AAAI Press, p. 138-148, 1997

UNESCO. **Guidelines for the establishment and development of monolingual thesauri**. Paris, 1973 *apud* CAMPOS, Maria Luiza de Almeida. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: UFF, 2001.

UNISIST. Princípios de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte. UFMG, Horizonte, v.10, n.1, p. 83-94, mar. 1981.

USCHOLD, M., KING, M. **Towards a Methodology for Building Ontologies**. 1995. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.55.5357>>. Acesso em: 30 jul. 2009.

VALE, E. A. do. Linguagens de indexação. In: SMITH, J. (Coord.). **Análise documentária: análise da síntese**. Brasília: IBICT, 1987. cap. 1, p. 12-26. p. 20.

VICKERY, B.C. Ontologies. **Journal of Information Science**, v. 23, n. 4, p. 277-286, 1997.

WANNER, L. Introduction. In: WANNER, L. **Lexical functions in lexicography and natural language processing**. Amsterdam: John Benjamins, 1996, p. 1-36.

WILLETS, Margaret. An investigation of the nature of the relation between terms in thesauri. **Journal of Documentation**, v. 31, n. 3, p. 158-184, 1975.

WORLD WIDE WEB CONSORTIUM. **OWL Web Ontology Language Overview:W3C recommendation**, 10 Feb. 2004. Disponível em: <<http://www.w3.org/TR/owl-features/>>. Acesso em: 30 jul. 2009.

WUESTER, E. L'étude scientifique générale de la terminologie, zone frontalière entre la Linguistique, la Logique, l'Ontologie, l'Informatique et les Sciences des Choses. In: RONDEAU, G.; FELBER, F. (Org.). **Textes choisis de terminologie: fondements théoriques de la terminologie**. Québec: GIRSTERM, 1981. p. 57-114 *apud* RIVIER, Alexis. **Construção de linguagens de indexação: aspectos teóricos**. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 21, n. 1, p. 56-99, jan./jun. 1992.

YU, C. T.; RAGHAVAN, V. V. Single-pass method for determining the semantic relationships between terms. **Journal of the American Society for Information Science**. V. 28, n. 6, p. 345-354, 1977.

ZUNIGA, G. **Ontology**: its transformation from philosophy to information systems. 2001. Disponível em: <<http://delivery.acm.org/10.1145/510000/505187/p187-zuniga.pdf?key1=505187&key2=1694381711&coll=&dl=acm&CFID=15151515&CFTOKEN=6184618>>. Acesso em 30 jul. 2009.