

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

TESAUROS E ONTOLOGIAS: ESTUDO DE DEFINIÇÕES
PRESENTES NA LITERATURA DAS ÁREAS DAS CIÊNCIAS DA
COMPUTAÇÃO E DA INFORMAÇÃO, UTILIZANDO-SE O MÉTODO
ANALÍTICO-SINTÉTICO

Orientadora
Lídia Alvarenga
Linha de Pesquisa
Organização e Uso da Informação

BELO HORIZONTE
2003

**TESAUROS E ONTOLOGIAS: ESTUDO DE DEFINIÇÕES
PRESENTES NA LITERATURA DAS ÁREAS DAS CIÊNCIAS DA
COMPUTAÇÃO E DA INFORMAÇÃO, UTILIZANDO-SE O MÉTODO
ANALÍTICO-SINTÉTICO**

ALEXANDRA MOREIRA

Dissertação apresentada à
Escola de Ciência da
Informação, da Universidade
Federal de Minas Gerais, como
requisito parcial para obtenção
do título de Mestre.

**BELO HORIZONTE
2003**

MOREIRA, Alexandra. Tesouros e Ontologias: *estudo de definições presentes na literatura das áreas das Ciências da Computação e da Informação, utilizando-se o método analítico-sintético*. Belo Horizonte, 2003. 150 f. il. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, 2003.

M838t

MOREIRA, Alexandra

Tesouros e Ontologias: estudo de definições presentes na literatura das áreas das Ciências da Computação e da Informação, utilizando-se o método analítico-sintético / Alexandra Moreira.

– Belo Horizonte, 2003.
150 f. il.

Orientador: Lídia Alvarenga.

Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, 2003.

1. Tesouros. 2. Ontologias. I. ALVARENGA, Lídia.
II. Universidade de Minas Gerais. Escola de Ciência da Informação.
III. Título.

CDD: 006.332
CDU:004.823:025.43

**TESAUROS E ONTOLOGIAS: ESTUDO DE DEFINIÇÕES
PRESENTES NA LITERATURA DAS ÁREAS DAS CIÊNCIAS DA
COMPUTAÇÃO E DA INFORMAÇÃO, UTILIZANDO-SE O MÉTODO
ANALÍTICO-SINTÉTICO**

ALEXANDRA MOREIRA

SINOPSE

Análise do termo ontologia usado na ciência da computação e do termo tesouro usado na ciência da informação com objetivo de se perceber as convergências destes dois instrumentos de representação do conhecimento. Para tanto foi aplicado o método analítico-sintético que permite o registro dos significados subjacentes a cada definição. A análise foi feita tomando-se como ponto de partida as definições retiradas da literatura que versam sobre ontologias e sobre tesouros. Estas definições foram rotadas no KWIC gerando os termos que foram analisados quantitativamente e qualitativamente. Conclui-se que tesouros e ontologias são instrumentos que operam no mesmo nível, o nível epistemológico, mas com propósitos e origem distintos.

**TESAUROS E ONTOLOGIAS: ESTUDO DE DEFINIÇÕES
PRESENTES NA LITERATURA DAS ÁREAS DAS CIÊNCIAS DA
COMPUTAÇÃO E DA INFORMAÇÃO, UTILIZANDO-SE O MÉTODO
ANALÍTICO-SINTÉTICO**

ALEXANDRA MOREIRA

ABSTRACT

This is a comparative analysis of the term ontology, used in the computer science domain, with the term thesaurus, used in the information science domain. The aim of the study is to establish the main convergence points between these two knowledge representation instruments and to point out their differences. In order to fulfill this goal an analytical-synthetic method was applied to extract the meaning underlining each one of the instruments selected definitions. The definitions were obtained from texts well accepted by the research community from both areas. The definitions were applied to a KWIC system in order to rotate the terms that latter were examined qualitatively and quantitatively. This research conclusion is that thesaurus and ontologies operate in the same knowledge level, the epistemological level, although with different origins and purposes.

AGRADECIMENTOS

À CAPES pelo o apoio financeiro aos meus estudos e pesquisa.

À Escola de Ciência da Informação da UFMG pela oportunidade.

À professora Lídia Alvarenga pela orientação, amizade e compreensão na realização deste trabalho.

Aos membros da banca pelas contribuições que serviram para aprimorar o texto final.

À todos os professores, colegas de curso e funcionários que ajudaram ao longo do curso com apoio, informações, esclarecimentos e idéias.

À professora Maria Luiza de Almeida Campos por ter me enviado a sua tese de Doutorado, cujo conteúdo muito contribuiu para este trabalho.

Ao meu marido que muito me ajudou no decurso do mestrado.

Aos meus pais e irmã pelo carinho e apoio.

Aos pais do meu marido pela compreensão.

À Nicollas Cuozzo, aluno de dança, que me incentivou a prestar o vestibular.

Ao meu grande e maravilhoso Deus, amigo inseparável, em todos os momentos.

SUMÁRIO

1	INTRODUÇÃO.....	10
2	DEFINIÇÃO DO PROBLEMA DE PESQUISA E JUSTIFICATIVA.....	12
3	OBJETIVOS GERAIS	15
3.1	OBJETIVOS ESPECÍFICOS.....	15
4	FUNDAMENTAÇÃO TEÓRICA.....	16
4.1	TEORIA DO CONCEITO.....	16
4.1.1	Conceitos, Características e Categorias	17
4.1.2	Intensão e Extensão do Conceito.....	20
4.1.3	Definição	21
4.2	TESAUROS E ONTOLOGIAS A PARTIR DA LITERATURA CONSULTADA: SUBSÍDIOS PARA A ABORDAGEM QUALITATIVA.....	23
4.2.1	Tesouros: gênese e definições.....	23
4.2.1.1	Função do tesouro	29
4.2.1.2	Tipos de tesouro	29
4.2.2	Ontologia.....	30
4.2.2.1	Ontologia na filosofia.....	31
4.2.2.1.1	A origem do termo.....	34
4.2.2.1.2	Ontologia na filosofia de Aristóteles	34
4.2.2.2	Influência da filosofia aristotélica no âmbito dos tesouros.....	37
4.2.2.3	Influência da filosofia aristotélica no âmbito da inteligência artificial	38
4.2.3	Ontologia na ciência da computação e na inteligência artificial	39
4.2.3.1	Origens e motivações.....	39
4.2.3.2	Ontologia como um sistema conceitual subjacente a uma base de conhecimento... 43	
4.2.3.3	Ontologia como um tipo especial de base de conhecimento.	44
4.2.3.4	Ontologia como um vocabulário usado por uma teoria lógica.....	46
4.2.3.5	Ontologia como uma especificação de uma conceitualização.	48
4.2.3.6	Linguagens para representação de ontologias	54
4.2.3.7	Considerações sobre a ontologia na ciência da computação e inteligência artificial 58	
5	METODOLOGIA.....	60
5.1	MÉTODO ANALÍTICO-SINTÉTICO.....	62
5.2	PROCESSO DE SELEÇÃO DOS TEXTOS	64
5.3	MATERIAL.....	68
5.4	O KWIC	69
5.5	ESCOLHA DOS TERMOS	70
6	RESULTADOS	72
6.1	CATEGORIAS	72
6.2	ATRIBUIÇÃO DOS TERMOS ÀS CATEGORIAS.....	73
6.3	TERMOS SELECIONADOS DAS DEFINIÇÕES SOBRE ONTOLOGIA CLASSIFICADOS POR CATEGORIAS	74
6.4	TERMOS SELECIONADOS DAS DEFINIÇÕES SOBRE TESAUROS CLASSIFICADOS POR CATEGORIAS	78
6.5	QUANTIFICAÇÃO DOS TERMOS	84
6.6	SISTEMA DE CONCEITOS.....	87
6.7	ANÁLISE SOBRE A OCORRÊNCIA DOS TERMOS NAS DEFINIÇÕES SOBRE ONTOLOGIAS	89
6.8	ANÁLISE SOBRE A OCORRÊNCIA DOS TERMOS NAS DEFINIÇÕES SOBRE TESAURO	91
6.9	ANÁLISE COMPARATIVA SOBRE A OCORRÊNCIA DOS TERMOS NAS DEFINIÇÕES SOBRE ONTOLOGIA E TESAURO.....	94

7	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS.....	96
8	ÍNDICE.....	101
9	BIBLIOGRAFIA.....	102
10	REFERÊNCIA BIBLIOGRÁFICA.....	110
11	ANEXOS.....	117
11.1	ANEXO I.....	117
11.2	ANEXO II.....	118
11.3	ANEXO III.....	119
11.4	ANEXO IV.....	129
11.5	ANEXO V.....	131
11.6	ANEXO VI.....	133
11.7	ANEXO VII.....	135
11.8	ANEXO VIII.....	137
11.9	ANEXO IX.....	147
11.10	ANEXO X.....	149

LISTA DE FIGURAS

Figura 1 - Triângulo do conceito	18
Figura 2 - Hierarquia de características	19
Figura 3 - Categorias	20
Figura 4 - Intensão e extensão de um conceito	21
Figura 5 - Evolução dos padrões de Tesouros	26
Figura 6 - Exemplo do relacionamento entre fatos e categorias	36
Figura 7 - Árvore de Brentano para as categorias de Aristóteles	36
Figura 8 - Correlação das categorias de Aristóteles com as categorias de Dahlberg	37
Figura 9 - Exemplo de base de conhecimento	45
Figura 10 - Especificação explícita de uma conceitualização	50
Figura 11 - Especificação explícita de uma conceitualização na forma gráfica	50
Figura 12 - Categorização de conceitos	51
Figura 13 - Níveis de representação para o conhecimento	52
Figura 14 - Exemplo de representação de ontologia por meio da linguagem OIL	56
Figura 15 - Tradução da ontologia para tesouro	57
Figura 16 - Trecho do resultado da consulta usando a palavra “ontologies”	66
Figura 17 - Número de citações em função do ano de publicação dos artigos.	66
Figura 18 - Exemplo do uso KWIC	70
Figura 19 - Sistema de conceitos para ontologia	88
Figura 20 - Sistema de conceitos para tesouros	89

LISTA DE TABELAS

TABELA 1 - Distribuição dos termos por categoria	85
TABELA 2 - Frequência de Ocorrência dos termos por categoria	87

LISTA DE GRÁFICOS

GRÁFICO 1 - Ocorrência de termos e categorias por definição de ontologia	90
GRÁFICO 2 - Ocorrências de termos de ontologia por categoria	91
GRÁFICO 3 - Ocorrência de termos e categorias por definição de tesouro	93
GRÁFICO 4 - Ocorrências de termos de tesouro por categoria	94

LISTA DE SIGLAS

AICHE - American Institute of Chemical Engineers
AFNOR - Association Française de Normation
ANSI - American National Standards Institute
BS - British Standards
CC – Ciência da Computação
CI – Ciência da Informação
COSATI - Committee on Scientific and Technical Information
DIN - Deutsches Institut für Normung
EJC - Engineers Joint Council
ISO - International Organization for Standardization
IA – Inteligência Artificial
TC – Teoria do Conceito
TEST Thesaurus of Engineering and Scientific Terms

LISTA DE SÍMBOLOS

\forall - Quantificador universal (para todo).
 \exists - Quantificador existencial (existe).
 \wedge - Conectivo lógico "e".
 \vee - Conectivo lógico "ou".
 \Rightarrow - Conectivo lógico para implicação.

1 INTRODUÇÃO

A escolha do tema do projeto de dissertação, *tesauros e ontologias*, foi motivada pela observação de que existe uma falta de entendimento sobre esses dois instrumentos de representação e recuperação de informação, cujas possíveis aplicações se sobrepõem, tornando difícil distingui-los a partir do seu uso.

Alguns cientistas da Computação encaram os tesauros como ontologias informais. Já na área da ciência da informação o termo ontologia é praticamente ignorado, sendo que apenas alguns poucos trabalhos mais recentes, de autores da área, mencionam esse instrumento (Campos, 2001b).

No entanto, a importância da ontologia como ferramenta de representação e recuperação da informação cresce a cada dia, tornando imprescindível o domínio de seu conceito por parte dos profissionais da ciência da informação. Sob a denominação de ontologias, têm sido criadas ferramentas que auxiliam a inserção e recuperação de documentos em sistemas computacionais (Guarino, 1999), a extração de informação em textos de linguagem natural, em sistemas de comércio eletrônico, a troca de informações entre agentes inteligentes¹ (Cranefield *et al.*, 2000), a aquisição automática de conhecimento (Duarte, 2002) e várias outras tarefas que envolvam o uso e a representação de conhecimento.

Este trabalho compreende um estudo da literatura sobre tesauros e ontologias, com o objetivo de contribuir com o esclarecimento terminológico e com o entendimento desses dois conceitos.

A apresentação desta dissertação obedece às seguintes etapas: no próximo capítulo é caracterizado o problema de pesquisa, mostrando a necessidade do esclarecimento das distinções e semelhanças entre os dois instrumentos; no capítulo 3 são apresentados os objetivos da pesquisa; no

¹ Agentes Inteligentes são programas de computador que simulam algum comportamento humano para resolução de tarefas. Em alguns sistemas os agentes precisam trocar informação para trabalhar cooperativamente para resolver um determinado problema.

capítulo 4 é apresentada a fundamentação teórica para o método de análise de conteúdo, como também, é apresentado um texto sobre tesauros e ontologias elaborado a partir da literatura consultada; no capítulo 5 é definida a metodologia de trabalho; no capítulo 6 são apresentados os resultados empíricos e a interpretação dos resultados e, finalmente, no capítulo 7 são apresentadas as conclusões.

2 DEFINIÇÃO DO PROBLEMA DE PESQUISA E JUSTIFICATIVA

Atualmente, bases de informação de todos os tipos têm surgido em profusão, com suporte de diferentes mídias. As iniciativas para disponibilizar essas bases de dados numa grande rede global de computadores também crescem acentuadamente. Com a popularização das ferramentas para armazenamento e divulgação em meios digitais, tornou-se impossível uma avaliação e uma estruturação adequadas das informações, sem o auxílio de métodos automáticos de classificação.

Para o processo de classificação e recuperação de informação, vários métodos automáticos vêm sendo desenvolvidos. No entanto, a maioria deles baseia-se em aspectos sintáticos e estatísticos, levando em consideração a frequência e distribuição de palavras presentes em documentos. Por se basear nos aspectos sintáticos da informação, a classificação realizada por esses métodos ainda está muito distante em termos de qualidade, da realizada por especialistas. Já a classificação automática de documentos baseada nos aspectos semânticos da linguagem, de modo geral ainda não é possível devido à complexidade de se obter a semântica de uma palavra, de uma unidade lexical, em um determinado contexto.

Outro problema é a troca de informação entre sistemas. Com a conexão entre sistemas de uma mesma empresa ou de empresas diferentes, as trocas informacionais precisam ser suportadas por meta-informações que indiquem o sentido das informações que estão sendo comunicadas. Caso contrário, não será possível garantir que as trocas informacionais sejam efetivadas sem inconsistências.

No contexto atual, existe uma grande demanda para desenvolvimento de sistemas que trabalhem com recuperação e troca de informação ou conhecimento. Para servir de suporte a esses sistemas e auxiliar a suprir essa demanda, novos instrumentos para a organização de conhecimento surgem a cada dia, sendo que, recentemente, muitos deles têm sido

denominado "ontologias". O emprego do termo ontologia para denominar uma estrutura de termos e as relações entre eles em um determinado domínio é mais comum na área da ciência da computação e, mais particularmente, na subárea da inteligência artificial. Alguns pesquisadores (Jasper e Mike Uschold, Fensel) consideram os tesouros como ontologias simples, uma vez que uma ontologia complexa, segundo esses autores, exige uma riqueza maior de relações do que as tradicionalmente apresentadas em um tesouro. Pode-se citar, como exemplo, a seguinte passagem de Fensel *et al.* (2001, p.): "*Larges ontologies such as WordNet provide a thesaurus for over 100,000 terms explained in natural language*". Neste sentido, pode-se entender os tesouros como sendo um tipo de ontologia voltada para organização de termos. Afirmarções com tal conteúdo denotam a existência de conceitos imprecisos de tesouros e ontologias. Daí a necessidade de um maior esclarecimento sobre esses instrumentos.

Seriam tesouros e ontologias termos que denotam o mesmo instrumento ou denotariam objetos diferentes e, portanto, seu uso indistinto seria inadequado? No segundo caso, estaria estabelecido um equívoco terminológico e as conseqüências dessa confusão seriam a falta de entendimento sobre o que caracteriza cada uma das ferramentas, resultando em possíveis inadequações nos processos de produção e uso. O desenvolvimento das ferramentas é especialmente prejudicado, uma vez que a confusão terminológica dificulta a troca de informações entre pesquisadores.

O uso do termo ontologia tornou-se muito freqüente na ciência da computação, principalmente na subárea da representação do conhecimento. Um dos principais objetivos do uso de ontologias na ciência da computação (CC) é a construção de bases de conhecimento interoperáveis e melhor estruturadas. A justificativa de um estudo para compreensão do termo dentro da área da ciência da informação (CI) decorre do fato da representação do conhecimento ser também objeto privilegiado de estudo no âmbito da ciência da informação. Por isso a contribuição das ontologias para a CI deve ser investigada. Além disso, um estudo de análise de conteúdo de um conceito

se encaixa no escopo da CI, especialmente quando se utiliza o método analítico-sintético, instrumento clássico de seu corpo teórico. A comparação do termo tesouro deve-se ao fato de alguns pesquisadores da CC defini-lo como uma “ontologia informal” e, portanto, trata-se de um aspecto que deve ser investigado.

3 OBJETIVOS GERAIS

- Buscar a compreensão semântica do termo ontologia à luz das diversas áreas onde é utilizado: ciência da computação, inteligência artificial e filosofia;
- Compreender o termo tesouro utilizado como ferramenta de representação de informação da ciência da informação;
- Contribuir para aproximar a ciência da informação da pesquisa sobre ontologias.

3.1 Objetivos específicos

Os objetivos específicos deste trabalho são:

1. Identificar e selecionar as definições de tesouros e ontologias;
2. aplicar o método analítico-sintético nos dois contextos;
3. comparar e refletir sobre ambos os conceitos.

4 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica que se segue possui duas funções. A primeira função é fornecer uma base teórica para a metodologia utilizada nesta pesquisa. Uma vez que o objetivo geral deste trabalho pode ser entendido como uma compreensão do significado de dois termos em relação a duas áreas de estudo, procurou-se utilizar uma metodologia que, partindo de definições selecionadas segundo certos critérios, fosse capaz de guiar a busca por um significado com ampla aceitação. A metodologia utilizada neste trabalho é uma adaptação da utilizada por Alvarenga (1993), que, por sua vez, possui raízes no método analítico-sintético que está inserido no contexto da Teoria do Conceito (TC), proposta por Dahlberg nos anos 70. Portanto, faz-se necessário uma apresentação da TC, bem como do método analítico-sintético, para um melhor entendimento da metodologia adotada. A TC será abordada neste capítulo, enquanto que o método analítico-sintético será abordado no capítulo sobre a metodologia. Como benefício adicional existe o fato da TC abordar o significado de termos tais como “conceitos”, “termos”, “categorias” e “definições”, que serão abordadas ao longo do trabalho e que são fundamentais nas discussões sobre tesouros e ontologias.

A segunda função relacionada com a fundamentação teórica é a de fornecer subsídios para uma discussão qualitativa sobre tesouros e ontologias. Para cumprir esta função foi feita uma pesquisa na literatura sobre ambos os termos objetivando fornecer uma visão da origem, evolução e características dos tesouros e das ontologias.

4.1 Teoria do Conceito

O estudo do conceito e suas relações é fundamental para a ciência da informação uma vez que permite fundamentar a análise de assunto para o processo classificatório, além de permitir o desenvolvimento de sistemas de

classificação coerentes que acompanhem a evolução do conhecimento. Sendo as modelagens conceituais processos de estruturação do conhecimento, suas aplicações ultrapassam as fronteiras do desenvolvimento de sistemas de classificação e possuem aplicações em outras áreas, como o desenvolvimento de aplicações de inteligência artificial e sistemas de informação em geral. Este trabalho adota as noções de “conceitos”, “termos”, “objetos”, “categorias” e “definições”, fixadas pela TC e que são descritas a seguir.

4.1.1 Conceitos, Características e Categorias

Para Dahlberg (1978, p. 143), um *conceito* é uma unidade de conhecimento, sendo o conhecimento a totalidade de enunciados verdadeiros sobre o mundo. Esta definição permite que o conceito seja algo palpável, armazenável e mensurável, o que não é verdade em outras definições que colocam o conceito em um plano subjetivo. Além disso, o conceito pode ser visto como:

“...uma tríade formada por (A) um referente (qualquer objeto material ou imaterial, atividade, propriedade, dimensão, tópico, fato, etc.), (B) os enunciados (predicações) verdadeiros e essenciais sobre um referente que estabelecem as características sobre o referente e (C) o termo, que é forma externa e comunicável do referente e suas características.”² (Dahlberg, 1981, p16)

Nota-se que na visão de Dahlberg o objeto, o termo e os enunciados são aspectos de uma mesma entidade: o conceito. Para ilustrar melhor a relação entre os elementos que compõem o conceito Dahlberg lança mão de um triângulo, denominado de *triângulo do conceito*, onde cada vértice é um elemento da tríade, como mostra a FIG. 1.

² Tradução da autora.

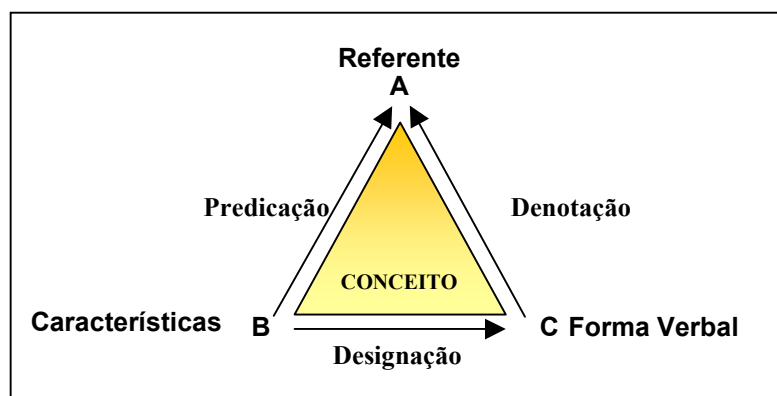


FIGURA 1 – Triângulo do conceito

Este triângulo difere do triângulo do significado, proposto por Ogden e Richards (1923), no sentido que no segundo o conceito é um dos vértices do triângulo, enquanto que, no triângulo proposto por Dahlberg, o conceito é o triângulo. Esta diferença reflete visões distintas sobre o relacionamento entre conceitos e termos. O triângulo do significado de Ogden e Richards reflete a visão de que não existe um significado "bem definido" associado a cada símbolo. Ou seja, adota a visão nominalista, que coloca o conceito em um nível mental e foca no símbolo. Esta é uma visão adequada aos lingüistas. Já o triângulo do conceito de Dahlberg adota uma visão conceitualista, que reúne as noções de conceito e termo como aspectos de uma mesma entidade. Esta é uma visão adequada à ciência da informação, uma vez que lida com linguagens especializadas, onde é preciso exigir uma correspondência unívoca entre termo e conceito.

Dahlberg distingue três tipos de conceitos: os *conceitos gerais*, os *conceitos especiais* e os *conceitos individuais*. Os conceitos gerais são aqueles que apontam para todos os itens de um dado tipo (Gênero – e.g. todos os periódicos semanais). Os conceitos especiais são aqueles que apontam para alguns itens de um dado tipo (Espécie – e.g. alguns periódicos semanais). E, finalmente, os conceitos individuais são aqueles que apontam para um único item de um dado tipo (Indivíduo – e.g. um periódico semanal).

Uma *característica* é o componente de um conceito que é derivado de

um enunciado sobre um item de referência e é um elemento do conteúdo de um conceito. Assim sendo, o conceito pode ser redefinido como uma unidade de conhecimento que engloba as características de um referente por meio de um termo ou nome.

Dahlberg distingue três tipos de características: *essenciais*, *acidentais* e *individualizantes*. As essenciais são características encontradas em todos os referentes de um dado conceito. As acidentais são encontradas em alguns referentes de um dado conceito. As individualizantes são encontradas em um único referente de um dado conceito.

As características podem arranjar-se de forma hierárquica, onde em cada nível são apresentadas características mais gerais, até que se atinja uma característica tão geral, que segundo Dahlberg pode ser considerada uma *categoria*. Ela define categoria como o conceito na sua mais ampla extensão. A FIG. 2, mostra a aplicação da hierarquia de características de um enunciado:

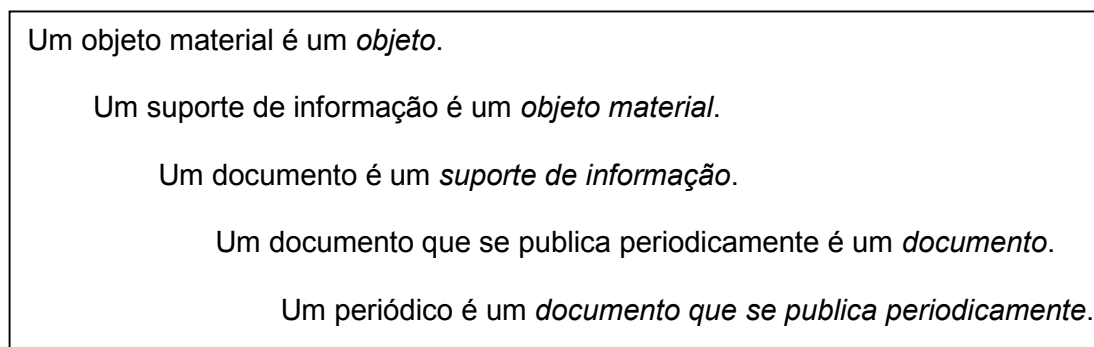


FIGURA 2 - Hierarquia de características
FONTE - DAHLBERG, 1978a.

Para Dahlberg as categorias possuem uma importância fundamental na organização dos conceitos. Elas permitem um entendimento sobre a natureza do conceito. Por exemplo, no caso do periódico a análise indicou que trata-se um objeto. Além disso, as categorias participam da organização da estrutura de conceitos. A análise do termo periódico em função de suas características gerou uma hierarquia que reflete a estrutura do conceito correspondente. As categorias de Dahlberg são baseadas nas categorias de

Aristóteles (mas não idênticas), e são compostas de quatro categorias que, por sua vez, são subdivididas em mais três subcategorias cada uma. A FIG. 3 mostra as categorias proposta por Dahlberg:

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

FIGURA 3 - Categorias
FONTE - DAHLBERG, 1978b, p.145.

4.1.2 Intensão e Extensão do Conceito

Dahlberg propõe a definição de intensão e extensão de um conceito a partir da noção de característica. Assim sendo a *intensão* de um conceito é definido como a soma total das suas características. Isto equivale à soma total dos conceitos mais genéricos mais as características específicas do conceito. Já a *extensão* de um conceito é definido como a soma total dos conceitos mais específicos e seus respectivos conceitos individuais. Pode também ser entendida como a soma total 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. A intensão e extensão de um conceito pode ser visualizada graficamente por meio da hierarquia de

características como a da FIG 4. Deste modo, a intensão de um conceito inclui os conceitos mais amplos, situados acima na hierarquia em relação ao conceito em estudo e a extensão engloba os conceitos mais específicos, situados na parte inferior da hierarquia, conforme ilustra a FIG. 4.

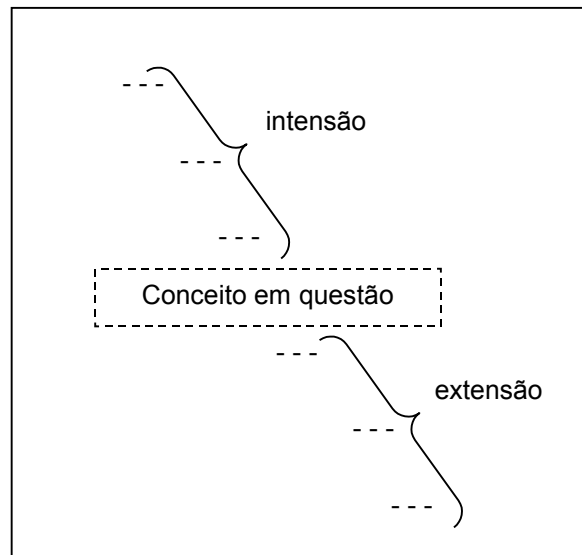


FIGURA 4 – Intensão e extensão de um conceito
FONTE - DAHLBERG, 1978b, p.148.

4.1.3 Definição

Com a definição busca-se a fixação precisa do significado de um conceito. Apenas após esta atribuição precisa do significado é possível usar os conceitos e seus termos de forma adequada. Consequentemente a obtenção e definição correta dos conceitos é fundamental, uma vez que o conhecimento se expande aceleradamente nos dias atuais conduzindo-nos à utilização de novos termos e conceitos cujo domínio nem sempre é fácil de manter.

Na definição trabalha-se com a intensão do conceito, uma vez que , se está lidando com suas características. Porém, no caso específico da definição nem todos os conceitos genéricos necessitam ser mencionados. Dahlberg (1978b) afirma que, “fazer uma definição equivale a estabelecer uma ‘equação de sentido’ sendo que, de um lado (à esquerda encontramos

aquilo que deve ser definido (o *definiendum*) e de outro (à direita) aquilo pelo qual alguma coisa é definida (o *definiens*)”. Pode-se compreender a definição como uma *limitação*, ou seja, uma colocação de limites sobre o entendimento de um conceito. Dahlberg ressalta que uma definição costuma ser escrita com o símbolo “df” para indicar expressamente a intenção de efetuar uma definição, como por exemplo:

morcego = df mamífero voador.

Existem vários tipos de definições, mas a distinção entre definições *nominais* e *reais* é a mais frequentemente abordada. Segundo Dalhberg (1978a) “a definição nominal tem por fim a fixação do sentido de uma palavra, enquanto que a definição real procura delimitar a intensão de determinado conceito distinguindo-o de outros com idênticas características”. Como exemplo de definição nominal temos:

opacidade = df não permeabilidade à luz.

Já a definição abaixo representa uma definição real:

bócio = df inchação do pescoço causada pela dilatação da glândula tiróide.

Assim, “a definição nominal relaciona-se com o conhecimento contido na linguagem” enquanto “a definição real relaciona-se com o conhecimento do objeto”, mencionando no *definiens* as características essenciais e, muitas vezes, as acidentais. Pode-se dizer que é uma tarefa difícil fazer a distinção entre a definição real e a nominal, sendo que, em alguns casos as duas ocorrem na mesma definição. Por exemplo:

discografia = df catalogação de discos (ou, descrição metódica dos discos de uma coleção. Cf. Aurélio)

4.2 Tesouros e ontologias a partir da literatura consultada: subsídios para a abordagem qualitativa

Como já foi mencionado, esta seção tem por objetivo fornecer uma visão da origem, evolução e características dos tesouros e das ontologias. Espera-se, assim, levantar fundamentos para ratificar a necessidade da pesquisa aqui proposta e para apoiar uma discussão envolvendo os dois termos, segundo uma ótica qualitativa. Um levantamento preliminar sobre a origem dos tesouros e das ontologias indicam alguns pontos de contato e diferentes motivações. No entanto, é interessante analisar o porque de alguns pesquisadores da ciência da computação insistirem em equipará-los. O levantamento a seguir pode fornecer algumas pistas para responder esta questão e para interpretar os resultados da análise de conteúdo das definições tal como proposto nesta dissertação.

4.2.1 Tesouros: gênese e definições

O termo “thesaurus” se origina do grego e do latim e significa “tesouro”. Este termo se popularizou com a publicação do dicionário analógico de Peter Mark Roget, em Londres, em 1852, intitulado “Thesaurus of English words and phrases”, que levou cerca de 50 anos para ficar pronto. Roget chamou de “thesaurus” seu dicionário de palavras, uma vez que o termo também designa vocabulário, dicionário ou léxico. O dicionário de Roget se diferenciava dos outros por ser um vocabulário organizado de acordo com o seu significado e não de acordo com a ordem alfabética. O tesouro de Roget teve o mérito de estabelecer a denominação para vocabulários que relacionam seus termos por meio de algum tipo de relação de significado; no entanto, alguns tesouros atuais pouco se assemelham com ele.

Nos anos 60, Vickery apresenta quatro significados para o termo tesouro na literatura da ciência da informação, sendo que o significado mais comum é o de uma lista alfabética de palavras, onde cada palavra é seguida de outras relacionadas a ela (Vickery *apud* Foskett, 1985, p.270). Howerton

apud Currás (1995, p. 85) define tesouro como “uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou intuitivas. Pode-se usar a lista manual ou mecanicamente, para indicar cabeçalhos de indexação”. Já Alan Gilchrist escreveu que tesouro é “uma lista autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação coordenada.”

Outro trabalho que deve ser destacado quando se fala de tesouro é o programa Unisist (UNESCO, 1973, p.6) que define “tesouro” para área da ciência da informação sob dois aspectos: o estrutural e o funcional. No primeiro caso, seria “um vocabulário controlado dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.” No segundo aspecto seria “um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita”.

Das primeiras iniciativas, como a de Mark Roget, até os dias atuais, os tesouros evoluíram em sua definição, construções teóricas e metodológicas, pela introdução de novos modelos cognitivos e abordagem centrada no usuário. Uma definição atual, resultante desta evolução é a de Currás (1995), que diz que

“Tesouro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que a compõem – termos, simples ou compostos – encontram-se relacionados entre si sintática e semanticamente.”

Em outras palavras, por linguagem especializada entende-se aquela que atua em um domínio restrito; por normalizada, compreende-se uma linguagem controlada; as unidades lingüísticas são termos e, finalmente, por pós-coordenada deseja-se indicar que os termos são combinados no momento de seu uso, em oposição às linguagens pré-coordenadas cujos termos que designam assuntos complexos se coordenam previamente à sua utilização (lista de cabeçalho de assunto, por exemplo).

Outra definição é a do *National Information Standards Organization*, no documento que propõe as linhas gerais para a construção, formatação e manutenção de tesouros monolíngues (ANSI/NISO Z39-19-1993), no qual tesouro é definido como um “vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão”.

Para o estudo da gênese e da linha evolutiva dos tesouros, utilizou-se o levantamento feito por Lancaster (1986). O autor ressalta a dificuldade de se mostrar esta evolução com precisão, uma vez que não está claro quem influenciou quem. A evolução histórica dos tesouros se apresenta sob duas vertentes: uma, que toma por base o Unitermo e a outra, influenciada pela Teoria da Classificação Facetada. Esta dicotomia é percebida em Lancaster, da seguinte forma: de um lado, a vertente da abordagem alfabética originária da América do Norte, mais especificamente os Estados Unidos, e do outro lado a influência da classificação bibliográfica da Europa, mais particularmente o Reino Unido. O quadro transcrito de Lancaster (1986) e exibido na FIG. 5 ilustra esse processo evolutivo.

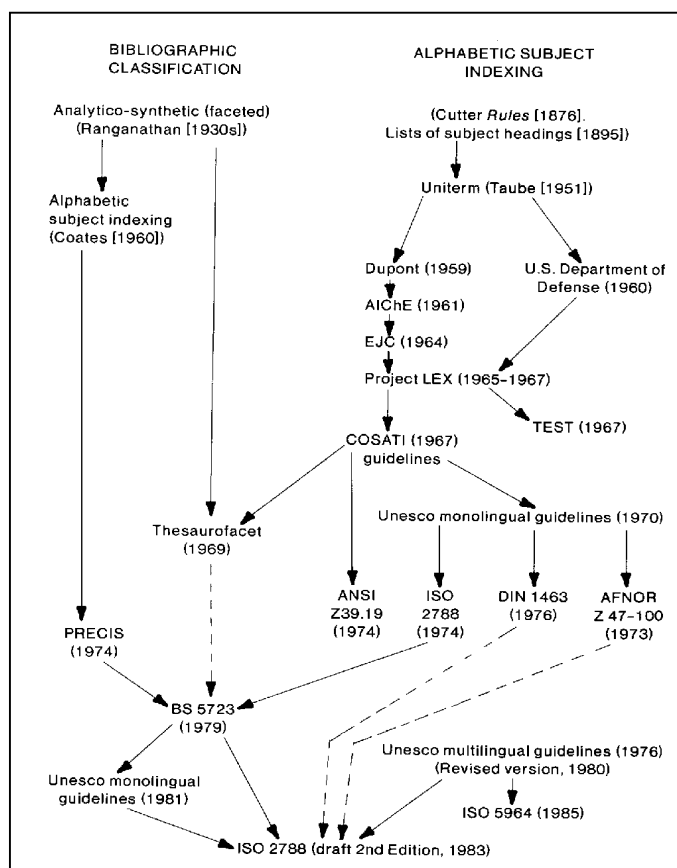


FIGURA 5 - Evolução dos padrões de Tesouros
 FONTE - LANCASTER, 1986, p. 30.

No quadro acima podem-se observar os trabalhos que influenciaram a possível evolução dos tesouros. De um lado, a Classificação Bibliográfica na Europa, iniciando com o trabalho de Ranganathan em 1930. O estudo analítico-sintético (facetado) de Ranganathan gerou princípios valiosos para esquemas de classificação, produzindo efeito na indexação alfabética de assunto e em tesouros (*Thesaurifacet*). Já a indexação alfabética de assunto, nos Estados Unidos, inicia-se com Charles Ammi Cutter, em seu *Rules for a Dictionary Catalogue* editado em 1876. Tanto a abordagem europeia quanto a outra convergem no padrão ISO 2788, texto de edição preliminar da segunda edição de 1983.

É importante ressaltar a importante contribuição de Mortimer Taube com seu sistema *Unitermo* em 1951 que, segundo Lancaster (1986), pode ser considerado responsável pelo aparecimento do tesouro. O *Unitermo* era composto por um conjunto de fichas, onde cada ficha continha uma única palavra e os números dos documentos associados a esta palavra. Segundo Gomes³, o *Unitermo* era fundamentado na hipótese de que cada idéia poderia ser representada por uma única palavra. Como pode ser visto na FIG. 5, a evolução do Unitermo levou à criação do primeiro tesouro, desenvolvido pelo Centro de Engenharia de Informação DuPont em 1959.

Pode-se perceber a participação e as contribuições na FIG.5 do AICHE (American Institute of Chemical Engineers) que publicou o *Chemical Engineering Thesaurus*, o EJC (Engineers Joint Council) em 1964 que publicou *Thesaurus of engineering Terms*, entre outros participantes. Dessa forma, estas colaborações serviram de base para novas tentativas, ou induziram a participação de outros interessados.

As duas linhas evolutivas de tesouros, europeia e americana, possuem algumas distinções importantes devido às motivações de seus desenvolvimentos. A americana adota uma evolução mais pragmática, motivada pela ampliação e pelo contorno das limitações de uma linguagem documentária preexistente, o *Unitermo*. No entanto, esse pragmatismo inibe o desenvolvimento de uma melhor fundamentação teórica para o instrumento. A consequência dessa ausência de um melhor embasamento teórico são as diferentes interpretações dos elementos dos tesouros. Gomes³ levanta alguns problemas dessa linha evolutiva, cujo instrumento é também denominado de “tesouro documentário”:

Assim, ora as relações partitivas eram consideradas como genérico-específicas, ora como associativas; também pela ausência de princípios freqüentemente fica-se sem entender porque dois termos estão ligados associativamente.

³ Artigo produzido por Gomes, H. E., disponível online no sitio www.conexaorio.com/bit/, intitulado, *Classificação, Tesouro e Terminologia: fundamentos comuns*.

Se havia mais de uma palavra para designar uma idéia - ou um tópico -, ora se considerava todas como parte integrante de um termo/descritor, ora não; ora uma expressão contendo mais de uma idéia era considerada como uma unidade (ou um termo), ora não. Com isto, as relações também não foram estabelecidas de forma clara.

Já a linha européia, principalmente a linha do *Thesaurofacet*, por se basear na teoria da classificação de Ranganathan (Lancaster, 1986, p.33), lança mão do uso de categorias para organização dos conceitos em um domínio. O uso de categorias para o enquadramento de conceitos permite uma melhor organização das hierarquias e um posicionamento mais adequados dos termos associados aos conceitos.

Como pode-se ver, ainda pela FIG. 5, existe uma tendência de junção das duas vertentes expressas pelo padrão BS 5723 (British Standards) e pela segunda edição da ISO 2788.

Outra linha mais recente, cuja menção é imprescindível, é a dos tesouros-com-base-em-conceitos, também denominado de tesouros terminológicos. Este tipo de tesouro originou-se na junção da Teoria do Conceito de Dahlberg, desenvolvida a partir da década de 70, com a Teoria da Classificação (Campos, 2001). Dentre as contribuições da Teoria do Conceito podem-se citar um melhor entendimento do conceito e do termo, a organização dos conceitos por meio de categorias, e o uso de definições para o posicionamento do conceito no sistema de conceitos.

Acredita-se que a discussão que esta proposta de pesquisa pretende realizar sobre tesouros e ontologias, será independente no que se refere à linha de evolução dos tesouros. Porém, acredita-se também, que as distinções apresentadas, entre as ontologias e os tesouros, ficarão mais evidentes no caso dos tesouros que utilizam categorias na organização de seus elementos.

4.2.1.1 Função do tesouro

Segundo Currás (1995), o tesouro foi adotado “na área de documentação, associado à forma de organização do vocabulário de indexação/recuperação”. O tesouro pode funcionar, num ambiente organizacional, na representação dos assuntos dos documentos e nas buscas informacionais. A representação dos assuntos dos documentos é realizada apenas pelo indexador que analisa o documento, identifica seu conteúdo e depois "traduz" para os termos permitidos de um tesouro.

O tesouro pode ser usado tanto para auxiliar na elaboração de consultas realizadas pelo usuário (consulente), em suas buscas informacionais, quanto pelo indexador durante o processo de classificação. Para os dois tipos de usuário, o tesouro, pela sua estrutura de termos e suas relações, auxilia a encontrar o melhor termo ou termos que representem um assunto. Portanto, o tesouro é um componente muito importante num sistema de recuperação por cumprir o papel de:

- determinar quais termos podem ser usados no sistema;
- determinar quais termos podem ser usados na busca para que esta tenha um resultado satisfatório; e
- permitir a introdução de novos termos em sua estrutura de termos e relações de modo a aproximar a linguagem do usuário à do sistema e realizar alterações de sentidos dos termos existentes.

4.2.1.2 Tipos de tesouro

A classificação de tesouros, aqui apresentada, baseia-se no Manual de Elaboração de Tesouros (Gomes, 1990) e em algumas outras bibliografias complementares. Os tesouros podem variar conforme a língua, nível de especificidade e assunto que cobrem.

a) Em relação à língua – podem-se distinguir os tesouros pelos idiomas em que se pode escrever seus termos em:

- **Monolíngues**
- **Multilíngues**

b) Em relação ao nível de especificidade:

- **Macrotesauros:** os termos representam conceitos mais ou menos amplos, resultando em um menor número de descritores. Como consequência, o número de remissivas é elevado. Já Currás declara que "um macrotesauro é composto de vários microtesauros, relacionados entre si por referências cruzadas, abrangendo uma ampla gama de assuntos, cada um deles constituindo objeto de microtesauro especializado". Um exemplo de tesouro clássico nesta categoria é o OCDE.
- **Microtesauros:** os descritores denotam conceitos em um nível maior de especificidade e se referem a um domínio mais restrito.

c) Em relação ao escopo ou assunto que cobrem:

- Projetados para atuar em um problema ou missão: por exemplo, Meio-Ambiente. Neste caso o tesouro deve ser necessariamente multidisciplinar, pois envolve descritores das várias áreas relacionadas com o problema.
- Voltados para um assunto: é o caso de tesouros voltados para uma única disciplina científica, como por exemplo a biologia.

4.2.2 Ontologia

O termo "ontologia" e termos relacionados como "relações ontológicas" e "comprometimentos ontológicos" surgem em diversas áreas do conhecimento, nem sempre sob a mesma interpretação. Esta seção visa mostrar as diversas visões sobre ontologia, na filosofia e na ciência da computação, particularmente na inteligência artificial. É apresentada,

também, uma discussão sobre as linguagens utilizadas na ciência da computação para a representação de ontologias e um pequeno exemplo de ontologia expressa em uma dessas linguagens.

4.2.2.1 Ontologia na filosofia

A ontologia é uma das divisões do estudo da filosofia. Segundo García Morente (1964), a filosofia possui duas grandes divisões: a ontologia (teoria do ser), relacionada com a filosofia da antigüidade e da idade média; e a gnosiologia (teoria do saber), relacionada com a idade moderna.

A ontologia, em linhas gerais, pode ser compreendida como a teoria do ser, mas, na visão de García Morente, uma significação rigorosa entende sua formação pelo genitivo **ontos**, que é o genitivo de **to on**; o genitivo **tou ontos** não significa ser, mas ente, no particípio presente. Portanto, ontologia tomada com esta visão seria a teoria do ente e não teoria do ser. Por meio deste exame filológico da palavra “ser” e do verbo “ser” busca-se um maior entendimento da ontologia, como em Aristóteles que em um de seus livros de Metafísica, começa dizendo “o ser se diz de muitas maneiras”. A ontologia será tanto a teoria do ente (tentativa de classificar os entes, tentativa de definir a estrutura de cada ente, tentativa de perceber cada tipo de ente) como também a teoria do ser em geral (daquilo que todos os entes têm de comum e daquilo que os classifica como entes). É a teoria dos objetos, das estruturas ônticas, o que há na vida. Estas coisas podem ser divididas em esferas de coisas reais, ideais, valores e a vida em sua totalidade. Alcança-se assim as quatro esferas fundamentais da ontologia.

Para Morente a ontologia se ocupa do ser em sua acepção mais vasta e ampla, e não de um ser concreto particular. Portanto a ontologia busca responder as seguintes perguntas “O que é o ser?” e “Quem é o ser?”. Morente acredita que a primeira pergunta “O que é o ser?” é impossível de ser respondida, uma vez que solicita como resposta uma definição do ser. Como toda definição pressupõe o enquadramento do objeto a ser definido em um conceito mais geral, é necessário enquadrar o conceito de ser em um

conceito mais geral. No entanto, não existe conceito mais geral do que o conceito de ser. A necessidade de enquadramento de um conceito em outro no que tange a definição de conceitos pode ser observada na seguinte menção de García Morente (1964, p. 58) :

...os conceitos de que nos valem nas ciências e até mesmo na vida, veremos que estes conceitos possuem todos eles uma determinada extensão, quer dizer que cobrem uma parte da realidade, que se aplicam a um grupo de objetos, a uns quantos seres. Mas estes conceitos são uns mais extensos que outros; quer dizer, que alguns se aplicam a menos seres que outros; como quando comparamos o conceito de “europeu” com o conceito de “homem”, encontramos, naturalmente, que há menos europeus do que homens. Por conseguinte, o conceito de “homem” se aplica a mais quantidade de seres que o conceito “europeu”. Os conceitos são, pois, uns mais extensos que os outros.

Ora: definir um conceito consiste em incluir este conceito em outro que seja mais extenso, e em outros vários que sejam mais extensos e que se encontrem, se toquem precisamente no ponto do conceito que queremos definir.

Ainda segundo Morente, no caso da segunda pergunta busca-se não a definição do ser mas apenas assinalá-lo e intuí-lo diretamente. Ou seja, procura-se dirigir a atenção do interlocutor para uma determinada situação ou local, onde está o conceito de ser. Uma vez assinalado, pode-se intuí-lo diretamente sem definição nenhuma. Morente diz que a pergunta “quem é o ser?” implica, também, na necessidade de distinguir o ser que é de verdade (o ser autêntico) do ser que não é de verdade (o ser falso ou inautêntico). Mas como realizar esta distinção? Para tanto deve-se perguntar: “que és?”. Se esta pergunta levar a dissolução ou a redução do pretendente a ser em outra coisa distinta então este não é um ser autêntico, mas um ser composto de outros seres.

Segundo Morente, a palavra “ser” pode ser vista em dois sentidos; de uma parte, existir, estar aí e o outro no sentido de consistir, ser isto, ser aquilo, ou seja, “qual a sua essência?”. Surgem, desta forma, novas indagações: “que é existir?”, “quem existe?”, “que é consistir?” e “quem consiste?”. Na opinião de Morente apenas as perguntas “quem existe?” e “que é consistir?” podem ser respondidas, sendo delas que a ontologia se

ocupa. A primeira pergunta não pode ser respondida porque a existência não pode ser definida mas apenas intuída por experiência íntima própria. Por razões semelhantes a pergunta “quem consiste?” não possui resposta uma vez que para isso é necessário responder, primeiramente, “quem existe?” e, de acordo com Morente, esta pergunta possui diversas respostas. As respostas para a pergunta “quem existe?” poderiam ser: eu existo, o mundo existe, Deus existe, as coisas existem; ou eu existo, porém as coisas não existem; ou apenas Deus existe e eu e as coisas existem em Deus; e assim por diante. O ramo da ontologia que se ocupa a responder esta pergunta é o denominado de Metafísica. A metafísica como mencionado em García Morente (1964, p.62)

é a parte da ontologia que se encaminha a decidir quem existe, ou seja, quem é o ser em si, o ser que não é em outro, que não é redutível a outro; e então os demais seres serão seres nesse ser em si. A metafísica é a parte da ontologia que responde ao problema da existência⁴.

Já a pergunta “que é consistir?”, “constitui a teoria do objeto, a teoria da objetividade ou a teoria da consistência dos objetos em geral” (Morente, 1964, p.62). Assim a ontologia vai tentar explicar o ponto de vista existencial metafísico e o ponto de vista objetivo, consistencial. E na busca de explicar “quem existe?” e “que é consistir?” o problema da ontologia vai convergir com o problema da metafísica que é “o que existe de verdade”. É neste momento que se atinge o fundamento de toda a realidade, a vida, esta sendo algo prévio e mais profundo do que a divisão do sujeito e do objeto, mas a tentativa de explicar a unidade que envolve estas quatro formas de objetividade (coisas reais, ideais, valores e a própria vida).

⁴ Grifo do autor do texto original.

4.2.2.1.1 A origem do termo

A palavra *ontologia* vem do grego *ontos* (ser) e *logos* (palavra). Apesar do estudo do ser ter suas raízes nos estudos de Aristóteles e Platão, o uso do termo ontologia para designar este ramo da filosofia é muito mais recente, tendo sido introduzido entre os séculos XVII e XVIII por filósofos alemães. Segundo Welty (2001), o termo foi cunhado em 1613 por Rudolf Goclenius e, aparentemente de forma independente por Jacob Lorhard. Segundo Mora (1963), o termo ontologia é mencionado brevemente por Goclenius na página 16 do *Lexicon philosophicum, quo tanquam clave philosophiae fores aperiuntur, Informatum opera studio Rodolphi Goclenii*, da seguinte forma: "ontologia, philosophia de ente", sendo esta a única menção ao termo. Coube a Christian Wolff, ainda segundo Mora, a popularização da palavra ontologia nos círculos filosóficos: "A palavra aparece no título de seu *Philosophia prima sive ontologia methodo scientifica pertractata, qua omnes cognitionis humanae principia continentur*, publicado em 1730."

4.2.2.1.2 Ontologia na filosofia de Aristóteles

O entendimento da visão de Aristóteles sobre ontologia merece uma discussão mais detalhada, devido à sua grande influência, tanto na ciência da computação, quanto na ciência da informação. Na ciência da informação, esta influência se faz notar nos fundamentos teóricos da área, como, por exemplo, o uso de categorias na teoria do conceito de Dahlberg e na classificação facetada de Ranganathan.

Aristóteles de Estagira nasceu em 384a.C. na cidade de Estagira que ficava em Calcídica. Por volta de 367 ou 366 a.C. chega em Atenas e se matricula na Academia de Platão freqüentando-a cerca de vinte anos. As obras de Aristóteles podem ser consideradas sob dois tipos de composição: as destinadas ao público em geral, sendo esta mais dialética; e as destinadas aos alunos do Liceu, sendo escritos filosóficos ou científicos. As obras de Aristóteles contemplam várias áreas científicas, mas boa parte do

material escrito se perdeu. Dentre as que restaram, a que está relacionada com o estudo presente é o tratado denominado de *Categorias*. Nela Aristóteles apresenta dez categorias básicas para classificar tudo o que existe. Ou seja, as categorias de Aristóteles revelam sua visão ontológica do mundo. Segundo García Morente (1964) as descrições das dez categorias propostas por Aristóteles são as seguintes:

- Substância - seria o que se pode dizer que algo "é" . Pode-se considerar a substância o que se diz de algo e isto é ele. Por exemplo: este é o homem, este é o gato, etc.
- Quantidade - pode-se dizer "quanto é", se é muito ou pouco. Por exemplo: um homem é pequeno ou grande, etc.
- Qualidade - pode-se dizer "que é". Por exemplo: é azul, bonito, etc.
- Relação - pode-se considerar uns seres em relação aos outros. Por exemplo: de um ser pode se predicar que é menor do que o outro, igual a outro, etc.
- Lugar - pode-se ante um ser determinar onde está. Por exemplo: está lá, em Paris, etc.
- Tempo - pode-se predicar "quando é", quando deixa de ser, quando foi. Por exemplo: século XX, é agora, etc.
- Ação - pode-se dizer "aquilo que aquele ser faz". Por exemplo: ato da semente que germinar.
- Paixão - pode-se dizer o que ele padece, sofre. Por exemplo: o gato é morto, etc.
- Posição - pode-se dizer sobre a posição do ser. Por exemplo: está em pé, sentado, etc.
- Estado - pode-se dizer sobre o estado do ser. Por exemplo: está florescido, seco, etc.

A primeira categoria é a substância que é a categoria mais abrangente e sobre a qual gira qualquer afirmação. As duas últimas categorias, "Posição"

e "Estado", não aparecem em algumas passagens, e parecem ter trazido dificuldades de caracter lógico e metafísico para Aristóteles. Para exemplificar o relacionamento entre os fatos de uma instância do mundo e as categorias aristotélicas pode-se usar o exemplo fornecido por Russell (2001), que descreve uma situação envolvendo Sócrates:

Substância	Sócrates
Qualidade	Filósofo
Quantidade	Cinco pés de altura
Relação	Amigo de Platão
Lugar	Na Ágora
Tempo	Ao meio-dia
Posição	De pé
Estado	Mal vestido
Ação	Conversando
Paixão	Sendo Insultado

FIGURA 6 – Exemplo do relacionamento entre fatos e categorias.
 FONTE - RUSSELL, 2001, p. 118.

Segundo Sowa (2000) o filósofo vienense Franz Brentano organizou as categorias de Aristóteles na forma de árvore, onde as categorias são os nós folhas e cujos ramos são rotulados por termos retirados de trabalhos de Aristóteles (FIG. 7).

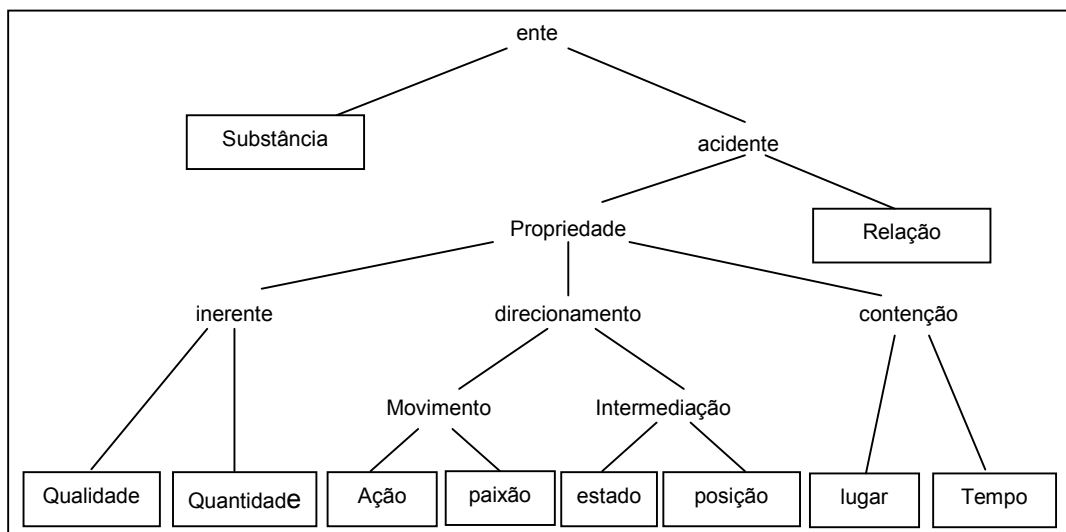


FIGURA 7 – Árvore de Brentano para as categorias de Aristóteles
 FONTE - SOWA, 2001.

4.2.2.2 Influência da filosofia aristotélica no âmbito dos tesouros

Como foi visto na seção sobre a gênese dos tesouros, estes evoluíram de listas de cabeçalhos de assunto até os tesouros baseados em conceitos. Os tesouros baseados em conceitos se fundamentam na Teoria do Conceito (TC). No preparo deste tipo de tesouro, após a seleção dos termos relevantes de um domínio utiliza-se como base as categorias formais da TC para iniciar a classificação dos termos. As categorias propostas pela TC baseiam-se nas fornecidas por Aristóteles. A FIG. 8, retirada de Campos (2001), mostra a correlação entre as categorias de Dahlberg e de Aristóteles.

Aristóteles		Categorias conceituais de Dahlberg
Substância	ENTIDADES	Fenômenos Objeto geral Objeto material
Quantidade Qualidade Relação	PROPRIEDADES	Contagem e mensuração Qualidade Comparação
Tempo Posição Espaço	DIMENSÕES	Tempo Posição Espaço

FIGURA 8 - Correlação das categorias de Aristóteles com as categorias de Dahlberg

FONTE - CAMPOS, 2001, p. 131.

A influência da filosofia de Aristóteles também pode ser observada no método de análise da TC. Aristóteles entendia que o homem conhecia os objetos e sua estrutura a partir das sensações. As sensações são o ponto de partida para que o homem reproduza como resultado da atividade intelectual, conceitos. Os conceitos seriam inerentes à estrutura mesma dos objetos e não idéias transcendentais do mundo físico. Assim o homem, por meio de observações que se repitam de casos particulares, poderia a partir

de uma operação intelectual, generalizar. Por exemplo: o conceito "gato" pode existir na mente humana como resultado, por via indutiva, da observação de vários seres concretos da mesma espécie. Aqui se percebe uma relação com o trabalho de Dahlberg que mostra também em sua teoria a possibilidade de se chegar a um entendimento ou uma classificação de um conceito a partir da indução e dedução.

4.2.2.3 Influência da filosofia aristotélica no âmbito da inteligência artificial

A contribuição mais óbvia da filosofia aristotélica para a IA é a lógica. A lógica é uma disciplina iniciada com Aristóteles com o propósito de garantir a veracidade das deduções. Até o século XVII a lógica não sofreu grande evolução, sendo que, somente com os trabalhos de Leibniz (1646-1716) e de Boole (1815-1864), a lógica foi modificada. O objetivo desses estudiosos era a formalização da dedução matemática, originando assim a Lógica Matemática. Outros pesquisadores contribuíram para a lógica matemática ter atingido a sua forma atual. Dentre eles pode-se destacar: G. Frege (1848-1925), B. Russell (1872-1970), N. Whitehead (1861-1947), Peano (1858-1932), D. Hilbert (1862-1943). Segundo Enderton (1972), a lógica matemática é "um modelo matemático do pensamento dedutivo". Já Quine, em seu livro *a filosofia da lógica* (1970), diz que a lógica é o "estudo sistemático das verdades lógicas".

O formalismo lógico é atraente porque sugere imediatamente uma maneira poderosa de derivar novos conhecimentos a partir dos antigos: a dedução matemática. Deste formalismo podemos concluir que uma nova frase é verdadeira, provando que ela é consequência de declarações já conhecidas. Assim, a idéia de uma prova, conforme desenvolvida na matemática, como uma maneira rigorosa de demonstrar a verdade de uma proposição em que já se acreditava, pode ser ampliada para incluir a dedução como uma maneira de derivar respostas a perguntas e soluções de

problemas. Devido a este fato a Lógica é utilizada por muitos pesquisadores de IA para a representação e dedução de conhecimentos.

Outra contribuição possível da filosofia aristotélica para a IA seria o uso das categorias para organização dos conceitos registrados nas bases de conhecimento, similarmente ao uso de categorias no desenvolvimento de tesouros. No entanto, a contribuição do aspecto ontológico somente nos últimos anos tem sido considerada e, mesmo assim, de forma diferente da indicada pela filosofia, como é ressaltado nos trabalhos de Valente (1996) e Guarino (1998).

4.2.3 Ontologia na ciência da computação e na inteligência artificial

Para se entender o uso do termo “ontologia” dentro da ciência da computação é preciso analisar suas origens na área e o que motivou seu surgimento. Além disso é preciso analisar os vários sentidos com que o termo vem sendo empregado no âmbito da ciência da computação, uma vez que o sentido do termo encontra-se em processo de amadurecimento.

4.2.3.1 Origens e motivações

O termo “ontologia” começou a ser empregado na ciência da computação, dentro da sub-área denominada de inteligência artificial (IA), no início dos anos 90, em projetos para organização de grandes bases de conhecimento, como CYC (Lenat e Guha, 1990) e Ontolingua (Gruber, 1992). Desde os anos 70 (Russel e Norvig, 1995) a IA já se preocupava com a organização e manipulação de bases de conhecimento, mas a partir da década de 90 houve um impulso para a criação de bases de conhecimento compartilháveis e reutilizáveis. Este impulso deveu-se à percepção de que os problemas complexos deveriam ser atacados por diferentes sistemas

atuando de forma cooperativa/competitiva em uma rede de multi-agentes⁵. Além disso, o esforço para criação de uma base de conhecimento pode ser muito caro e a reutilização e compartilhamento das bases podem baratear os custos. No entanto, o compartilhamento de bases de conhecimento só pode ocorrer se houver um claro entendimento dos “*comprometimentos ontológicos*” associados às bases. Por comprometimentos ontológicos entende-se as escolhas que levaram a selecionar um determinado conjunto de conceitos em vez de outro (Valente, 1995, p.34). Ou seja, os comprometimentos ontológicos determinam o que é relevante em um determinado domínio para que seja representado em uma base de conhecimento. Por exemplo, ao se escolher representar o objeto “livro” por meio de um predicado em uma linguagem lógica, o desenvolvedor está se “comprometendo” com a existência desta propriedade no domínio, ou seja, que existem objetos que possuem a propriedade de ser livro. A própria escolha da linguagem de representação revela alguns comprometimentos ontológicos. Por exemplo, o uso da lógica de primeira ordem revela o comprometimento ontológico com a existência de fatos, objetos e relações (Russel, 1995, p.166). O compartilhamento dos mesmos comprometimentos ontológicos é que torna possível a comunicação entre dois agentes (humanos ou não), uma vez que estabelece que os mesmos objetos foram abstraídos pelos agentes envolvidos.

O registro explícito e formal dos comprometimentos ontológicos é o que tem sido denominado, na maioria das vezes, de ontologia no âmbito da inteligência artificial. Enquanto as bases de conhecimento tradicionais acumulavam o conhecimento necessário para atender uma aplicação específica, uma ontologia deveria ter a propriedade de ser usada em várias aplicações e em aplicações distribuídas, como sistemas multi-agentes, fornecendo o suporte necessário para a troca de informação entre agentes.

⁵ Agentes na Inteligência Artificial é uma forma de classificar os componentes de um sistema que reagem às percepções recebidas do ambiente e que possuem autonomia e objetivos próprios. Dentro desta visão, várias entidades, naturais ou artificiais podem ser classificadas como agentes, indo desde certos tipos de programas até pessoas.

Jasper e Uschold (1999, p. 6) identificaram quatro categorias principais de aplicações para ontologias:

Autoria neutra: *Um artefato de informação é criado em um único idioma, e é convertido em uma forma diferente para uso em múltiplos sistemas alvo. Benefícios desta aproximação incluem reuso de conhecimento, melhoria da manutenção e retenção do conhecimento.*

Ontologia como especificação: *Uma ontologia sobre determinado domínio é criada e usada como base para especificação e desenvolvimento de software. Benefícios desta abordagem incluem documentação, manutenção, confiabilidade e reuso do conhecimento.*

Acesso comum à Informação: *A informação é requerida por uma ou mais pessoas ou aplicações de computador, mas é expressa em vocabulário pouco conhecido, ou em um formato inacessível. A ontologia torna a informação inteligível provendo um entendimento compartilhado dos termos, ou mapeando entre conjunto de termos. Benefícios desta aproximação incluem interoperabilidade, e uso e reuso mais efetivo de recursos de conhecimento.*

Busca baseada em ontologia: *Uma ontologia é usada para procurar um repositório de informação para recursos desejados (por exemplo documentos, páginas da Web, nomes de especialistas). O benefício principal desta abordagem é o acesso mais rápido a recursos de informação importantes que conduzem a um uso e reuso mais efetivo de recursos de conhecimento.*

Aqui discorda-se de Jasper e Uschold em relação ao benefício apontado por eles no que diz respeito ao uso de ontologia na "Busca baseada em ontologia". Neste caso, o maior benefício não é o acesso mais rápido a recursos de informação e sim um aumento da qualidade da informação retornada, uma vez que documentos fora do contexto têm maior chance de serem rejeitados. Dentre os problemas específicos em que o uso de ontologias pode contribuir, podem-se citar: representação, reuso, compartilhamento, aquisição e integração de conhecimento; processamento de linguagem natural; tradução automática; comunicação de informação entre sistemas, agentes, empresas ou pessoas.

Contudo, ainda não existe um consenso sobre esta interpretação do termo. Mesmo quando existe um acordo de que a ontologia para a IA é o registro dos comprometimentos ontológicos, a forma como é feito este registro é assunto para debate. Dentre os pesquisadores que tentaram

elucidar a interpretação do termo para a inteligência artificial destacam-se Guarino e Giaretta (Guarino E Giaretta, 1995) e Poli (Poli, 2001). Guarino e Giaretta (1995) apresentam sete interpretações para o termo ontologia, existentes na área da inteligência artificial e da representação do conhecimento. Já Poli analisa algumas das definições existentes do âmbito da IA segundo os critérios da orientação (ontologias orientadas a objeto e orientadas a conceitos) e da independência de domínio. Segundo ele, ontologias no sentido correto da palavra seriam orientadas a objetos e independentes de domínio, sendo que ontologias orientadas a conceitos e dependentes de domínio seriam as mais espúrias.

Para fins deste estudo serão discutidas, com maior detalhe, algumas das interpretações mais comuns para o termo ontologia, segundo as definições obtidas dos textos selecionados. As interpretações que serão analisadas são:

- 1. Ontologia como um sistema conceitual subjacente a uma base de conhecimento.*
- 2. Ontologia como um tipo especial de base de conhecimento⁶.*
- 3. Ontologia como um vocabulário usado por uma teoria lógica.*
- 4. Ontologia como uma especificação de uma conceitualização.*

As interpretações acima são um subconjunto das apresentadas por Guarino e Giaretta. A primeira interpretação considera a ontologia como um sistema conceitual que pode estar subjacente a uma base de conhecimento. Neste caso a ontologia é um objeto localizado no nível conceitual e não no nível simbólico. A partir do item 2 o termo ontologia é interpretada como denotando uma entidade no nível simbólico. A segunda interpretação define a ontologia como um tipo especial de base de conhecimento, diferenciando desta por conter um tipo especial de conhecimento (conhecimento independente de uma configuração especial do domínio) ou por servir a um

⁶ Usa-se o termo “base de conhecimento” no sentido de um conjunto de sentenças descrevendo o estado de um domínio na forma de uma teoria lógica.

propósito específico (comunicação, consulta, etc.). Segundo a interpretação 3, uma ontologia é apenas o vocabulário usado em uma teoria lógica, sendo que o nível de formalização na definição deste vocabulário pode ser bastante amplo. A interpretação 4 estabelece que a ontologia é uma especificação de uma conceitualização, sendo que uma conceitualização pode ser entendida como o conjunto dos comprometimentos ontológicos. Esta é a interpretação que está angariando o maior número de adeptos na IA. Estas interpretações serão discutidas com maior detalhamento a seguir.

4.2.3.2 Ontologia como um sistema conceitual subjacente a uma base de conhecimento.

Alguns pesquisadores não enquadram as ontologias como sendo artefatos concretos e sim como uma estrutura conceitual subjacente a uma base de conhecimento. Sowa (1995) descreve uma ontologia como sendo o conteúdo de formas lógicas. Em outras palavras, a ontologia seria a conceitualização subjacente à uma base de conhecimento, e, desta forma, uma entidade de nível conceitual. O surgimento da ontologia precede a criação da base de conhecimento, uma vez que envolve a escolha dos conceitos e relações que serão representados na base de conhecimento, escolhas essas relacionadas com os comprometimentos ontológicos pretendidos. Duas bases de conhecimento podem possuir a mesma conceitualização. Da mesma forma, uma vez que a ontologia pertence ao nível conceitual e não é apresentada de forma explícita no nível sintático, é possível que as sentenças em uma base de conhecimento sejam sujeitas a diferentes interpretações de comprometimentos ontológicos. Por exemplo, suponha que a partir da observação de que uma pessoa, cujo nome é Pedro possui um livro intitulado “O que é ontologia”; pode-se representar este fato por meio da seguinte sentença lógica:

$$\text{Pessoa(Pedro)} \wedge \text{Livro(O-que-é-ontologia)} \wedge \text{Possui(Pedro, O-que-é-ontologia)}$$

Subjacente a esta sentença existe o entendimento do predicado “Pessoa” representando o conceito de um indivíduo da espécie humana; o entendimento do predicado “Livro” representando o conceito de um objeto particular inanimado, etc.; e o entendimento da relação “Possui” representando o conceito de relação de posse de um objeto particular por um indivíduo da espécie humana. No entanto, uma vez que estes, e outros comprometimentos ontológicos não descritos aqui, não foram explicitados, é possível relacionar a sentença acima com a seguinte conceitualização: o predicado “Pessoa” representando o conceito de um indivíduo da espécie humana; o predicado “Livro” representando o conceito de uma obra intelectual como objeto abstrato (como na sentença “*fulano* escreveu o livro *tal*”); e o entendimento da relação “Possui” representando o conceito de relação de criação de uma obra por um indivíduo da espécie humana. Pode-se verificar, portanto, que ausência do registro dos comprometimentos ontológicos pode dificultar o entendimento do significado pretendido, uma vez que o simples uso de termos em linguagem natural como predicados e constantes é insuficiente para capturar estes compromissos. Pode-se alegar que a sentença não está corretamente representada e, por isso, existe a possibilidade de um entendimento diferente do pretendido. No entanto, a questão é que a falta do registro dos comprometimentos ontológicos aumenta a possibilidade de confusão.

Esta interpretação da ontologia como um sistema conceitual subjacente a uma base de conhecimento não é muito popular na ciência da computação, uma vez que, como não está situada no nível simbólico, não pode ser armazenada e operada computacionalmente.

4.2.3.3 Ontologia como um tipo especial de base de conhecimento.

Para alguns pesquisadores uma ontologia é uma base de conhecimento que se distingue das demais por: 1) possuir apenas um tipo determinado de conhecimento; ou 2) ser orientada a determinado tipo de

tarefa. Em ambos os casos, uma ontologia, é um artefato concreto no nível simbólico e, portanto, pode ser compartilhada e transmitida.

No primeiro caso, a maioria dos autores entende que o conhecimento registrado em uma ontologia deve descrever objetos e relações que estejam sempre presentes no domínio, independentemente do estado particular de um domínio. Lethbridge (1994) considera uma ontologia como um "tipo especializado de base de conhecimento que consiste de descrições ou definições dos tipos de coisas." O problema, com esta interpretação é que ela não determina que tipo de conhecimento deve ser representado em uma ontologia. Estabelecer que o conhecimento deve ser independente de um estado particular do domínio é muito vago. Por exemplo, seja a seguinte base de conhecimento:

<i>Sentença em Lógica</i>	<i>Linguagem natural</i>
$\forall x \text{ homem}(x) \Rightarrow \text{mortal}(x)$ homem(João)	Todo homem é mortal João é um homem

FIGURA 9 – Exemplo de base de conhecimento

A primeira sentença é sempre verdadeira, independente de um estado particular, ao passo que a segunda é verdade em algumas instâncias. Portanto, uma ontologia deveria conter a primeira sentença enquanto que a segunda ficaria de fora. De fato, a primeira sentença revela alguma informação sobre os conceitos envolvidos. Ela indica que a extensão do conceito de “homem” está contida na extensão do conceito “mortal”. É uma informação de subjugamento, ou taxionômica. No entanto, não revela maiores informações sobre a natureza dos conceitos envolvidos. O que é ser homem? É no sentido de ser humano ou no sentido de ser um espécime humano do sexo masculino? A ontologia deve registrar informações sobre a natureza dos conceitos de modo a restringir as interpretações possíveis? Muitos pesquisadores de IA acreditam que sim e por isso adotam a interpretação 4 que é mais específica que a corrente.

4.2.3.4 Ontologia como um vocabulário usado por uma teoria lógica.

Tanto a interpretação 2 quanto a 3 implicam em uma representação da ontologia na forma de uma teoria formal. Contudo alguns pesquisadores classificam a ontologia como um artefato sintático, mas não exigem que ela tenha o rigor de uma teoria formal. Alguns definem ontologia como sendo apenas o vocabulário adotado em um domínio específico. Por exemplo, Russell e Norvig (1995, p.222) relacionam a ontologia de domínio com bases de conhecimento, sendo a ontologia do domínio o vocabulário de predicados, funções e constantes que será utilizado na construção da base de conhecimento⁷. Para eles, a ontologia é uma lista informal de conceitos de um domínio. A definição dos elementos do vocabulário na forma de axiomas, em uma linguagem formal resulta em uma base de conhecimento. Na visão de Russell e Norvig, o produto final é a base de conhecimento, e a ontologia é um produto intermediário associado à base. Swartout *et al.* (1996, p.1) compartilham da mesma visão: “Uma ontologia é um conjunto de termos estruturados hierarquicamente para descrever um domínio que pode ser usado como um esquema base para uma base de conhecimento”. Já Chandrasekaran *et al.* (1999) fornecem uma definição mais completa:

Em um sentido uma Ontologia é um vocabulário de representação, frequentemente especializado para algum domínio ou assunto. Em um segundo sentido uma ontologia é usada para referir a um corpo de conhecimento descrevendo algum domínio, tipicamente um conhecimento comum de um domínio, usando um vocabulário de representação. (p. 1)

Jasper e Uschold (1999) incluem em sua definição a necessidade da existência das definições dos termos e seus relacionamentos:

⁷ Uma constante é um nome de um objeto (Pedro). Um predicado representa uma propriedade (ex.: Pedro é pai) ou uma relação (ex.: Pedro é pai de Ana) enquanto uma função representa um indivíduo (ex.: o pai da Ana). Note que o termo “pai” foi usado em todos os exemplos, mostrando que o desenvolvedor deve escolher a melhor forma para expressar os conceitos de um domínio. A escolha determina algumas perdas ou ganhos. Uma constante isolada não carrega nenhum significado. No caso da propriedade sabe-se apenas que Pedro é pai mas não sabe-se de quem. Já no caso da relação sabe-se que Ana é a filha. Finalmente, no caso da função é possível falar de um objeto sem saber seu nome (o pai da Ana comprou um carro).

Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos.(p. 2)

Dentre outros pesquisadores com definições semelhantes pode-se citar, Wiederhold (1994, p.6): “Uma ontologia é um vocabulário de termos e uma especificação de suas relações”; e Hovy (1998, p. 2): “Uma ontologia é um conjunto de termos, associado com definições em linguagem natural e, se possível, usando relações formais e restrições, sobre um domínio de interesse”.

Quando existe a exigência do uso de definições formais dos termos e de suas relações a interpretação corrente coincide com a interpretação 4, uma vez os termos denotam conceitos e o registro das definições e relações na forma de uma teoria formal pode ser vista como uma especificação de uma conceitualização.

Quando não existe exigência da representação das definições e das relações como uma teoria formal a interpretação corrente possibilita visualizar a ontologia como um tipo de tesouro. Alguns pesquisadores (Jasper e Uschold 1999, Fensel 2001) consideram os tesouros como ontologias simples, uma vez que uma ontologia complexa exige uma riqueza maior de relações. Como exemplo, repete-se aqui a seguinte passagem de Fensel *et al.* (2001): “*Large ontologies such as WordNet provide a thesaurus for over 100,000 terms explained in natural language*”. Neste sentido, pode-se entender os tesouros como sendo um tipo de ontologia voltada para organização de termos. Wordnet (Miller, 1995) é um banco de dados linguístico contendo termos agrupados em conjuntos de equivalência semântica. Os conjuntos são relacionados por vários tipos de relações semânticas, dentre as quais hiponímia, hiperonímia e antonímia. Sowa (2000) também considera o Wordnet como um tipo de ontologia, denominada de *ontologia terminológica*. Em um de seus artigos, Sowa

(2000) define ontologia terminológica como “uma ontologia cujas categorias não precisam ser totalmente especificadas por axiomas e especificações”, e acrescenta:

“Um exemplo de ontologia terminológica é o Wordnet, cujas categorias são parcialmente especificadas por relações tais como subtipo/supertipo ou parte/todo, que determinam a posição relativa dos conceitos em relação aos outros mas que não os definem completamente.” (Sowa, 2000)

Note que o comentário acima também pode ser feito em relação a um tesaurus típico. Ainda segundo Sowa, a diferença entre uma ontologia terminológica e uma terminologia formal é mais de grau do que de conteúdo: quanto mais axiomas são adicionados a uma ontologia terminológica mais próxima ela fica de uma ontologia formal ou axiomatizada.

Decerto que tesaurus são instrumentos terminológicos e que, obviamente, registram parcialmente os comprometimentos ontológicos, no entanto, enquadrá-los como um tipo de ontologia pode ser precipitado, uma vez que existem diferentes princípios envolvendo a construção de cada instrumento. Por exemplo, a construção de um tesouro envolve os princípios da garantia de uso e da garantia literária. Além disso, os tesaurus são desenvolvidos tendo em mente uma tarefa específica, enquanto que as ontologias deveriam ser, a princípio, independente de tarefas. Keizer *et al.* (2000), também não enquadram tesaurus como um tipo de ontologia. Eles preferem enquadrar tesaurus, classificações, vocabulários, nomenclaturas e sistemas de codificação sob a denominação de *sistemas terminológicos*. Segundo ele, o foco das ontologias é mais direcionado para os conceitos, enquanto que, nos sistemas terminológicos, o foco é mais direcionado para os termos.

4.2.3.5 Ontologia como uma especificação de uma conceitualização.

A interpretação 4 é a mais popular dentro da comunidade de representação de conhecimento e tem como representante mais conhecida

a definição proposta por Grubber (1993) que diz que uma ontologia é "*especificação formal e explícita de uma conceitualização compartilhada*", onde *formal* significa que a ontologia pode ser expressa em uma linguagem formal, *explícita* significa que é um objeto de nível simbólico e *compartilhada* indica que o conhecimento é aceito dentro da comunidade. Já uma *conceitualização* é, segundo Gruber,

"... os objetos, conceitos, e outras entidades que se assume existir em alguma área de interesse e as relações que ocorrem entre eles. Uma conceitualização é uma visão abstrata e simplificada do mundo que nós desejamos representar para algum propósito. Toda base de conhecimento, sistema baseado em conhecimento, ou agente atuando no nível do conhecimento é comprometido com alguma conceitualização, explícita ou implicitamente". (tradução da autora) (Grubber, 1993, p.1)

A definição de Grubber é a que tem sido mais adotada pelos pesquisadores da comunidade de representação do conhecimento, podendo ser vista nos trabalhos de Fensel *et al.* (2001), Uschold (1996a), Pontes (2002), Lethbridge (1994) Erdmann e Studer (1999) Van Heijst *et al.* (1997) e Duineveld (1999), somente para citar alguns. Se a ontologia for representada em uma linguagem formal a ontologia será denominada de *ontologia formal*. Uma definição semelhante pode ser encontrada Gruninger e Fox (1995, p.1) que diz que "uma ontologia é uma descrição formal de entidades e suas propriedades, relações, restrições e comportamentos". Neste caso, deve-se interpretar a "descrição formal de entidades e suas propriedades, relações, restrições e comportamentos" como uma "especificação formal e explícita de uma conceitualização", e que também coincide com a interpretação de conceitualização proposta por Gruber.

Um exemplo pode ser útil para ilustrar as noções de conceitualização e especificação. Suponha a visão do mundo formada pelo domínio acadêmico. Uma conceitualização desta visão poderia conter os conceitos de "aluno", "professor", "disciplina", "nota", "curso", "ministra", etc. Uma especificação explícita desta conceitualização em lógica de primeira ordem poderia ser algo como:

<i>Sentença em Lógica</i>	<i>Linguagem natural</i>
$\forall x \text{ disciplina}(x) \Rightarrow \exists y(\text{professor}(y) \wedge \text{ministra}(x, y))$	Para toda disciplina existe um professor que a ministra.
$\forall x \text{ disciplina}(x) \Rightarrow \exists y(\text{curso}(y) \wedge \text{parte-de}(x, y))$	Toda disciplina é parte de um curso.
$\forall x \text{ aluno}(x) \Rightarrow \exists y(\text{curso}(y) \wedge \text{cursa}(x, y))$	Todo aluno cursa um curso.
$\forall x \text{ aluno}(x) \Rightarrow \text{Corpo-acadêmico}(x)$	Todo aluno é um elemento do corpo acadêmico.
$\forall x \text{ professor}(x) \Rightarrow \text{Corpo-acadêmico}(x)$	Todo professor é um elemento do corpo acadêmico.
...	...

FIGURA 10 – Especificação explícita de uma conceitualização

Note que as descrições de um estado particular do domínio não são incluídas, tais como “João cursa Medicina” ou “Ana é professora”. Uma apresentação menos formal da conceitualização acima poderia ser feita usando uma notação gráfica, onde os nós são os objetos e os arcos são relações:

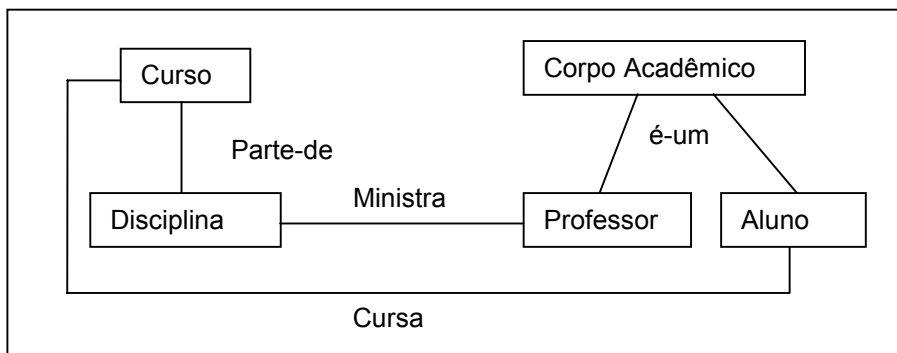


FIGURA 11 – Especificação explícita de uma conceitualização na forma gráfica

Apesar das representações acima capturarem uma porção de uma conceitualização, será que é este o conhecimento registrado por uma ontologia? Tradicionalmente, as representações acima seriam chamadas de bases de conhecimento e não de ontologia. Seria ontologia um novo nome para algo que já existia? Para Guarino (1995b) o conhecimento acima não é o conhecimento que deve ser registrado em uma ontologia. Segundo ele esse conhecimento pertence aos níveis denominados de *lógico* e *epistemológico* e não ao *nível ontológico*. No nível lógico o conhecimento é codificado por meio de primitivas básicas como predicados e funções. Já na

opinião de Guarino (1995b), a distinção entre a visão epistemológica e a visão ontológica na representação do conhecimento está no fato de que a primeira se preocupa com a *forma do conhecimento* e a última se preocupa com a *forma do ser*. Então, a epistemologia fornece uma estrutura para o conhecimento e a ontologia fornece uma descrição das formas de seus elementos. Assim sendo, os conceitos apresentados nos outros níveis são vistos como objetos e sua natureza, como tal, é observada. O registro desta observação é o que deve conter uma ontologia formal. Por exemplo, usando-se as categorias elaboradas por Dahlberg (1978b) pode-se registrar o seguinte conhecimento de natureza ontológica sobre os conceitos capturados do domínio acadêmico:

Todo Professor e aluno é um objeto material . Curso e disciplina são objetos imateriais . Cursar é um processo
--

FIGURA 12 – Categorização de conceitos

Pode-se observar que o conhecimento ontológico procura descrever a natureza dos objetos de forma independente de um problema específico. Uma vez que o propósito da modelagem é construir uma visão do mundo para comunicação de conhecimento, e não para a resolução de um problema específico, a necessidade do estabelecimento de compromissos ontológicos é mais importante do que a visão epistemológica. Portanto, a visão ontológica deve prevalecer sobre a visão epistemológica quando se trata de sistemas de recuperação de informação e compartilhamento de conhecimento. Para situar o papel da ontologia na representação do conhecimento, pode-se basear nos níveis de representação mostrados na figura a seguir. A FIG. 13 é uma adaptação da apresentada em Guarino (1995b). Estes níveis foram propostos, originalmente, por Brachman (1979) *apud* Guarino (1994). No entanto, a proposta de Brachman não incluía o nível ontológico.

<i>Nível</i>	<i>Primitivas</i>	<i>Exemplo</i>
Lógico	Predicados, funções	$\forall x \text{ aluno}(x) \Rightarrow \text{Corpo-acadêmico}(x)$ $\exists x \text{ aluno}(x) \wedge \text{Inteligente}(x)$
Epistemológico	Relações de estruturação	aluno é uma subclasse do corpo acadêmico. Existem alunos que são inteligentes.
Ontológico	relações Ontológicas	Todo aluno é um objeto material . Inteligente é uma qualidade .
Conceitual	Relações Conceituais	
Lingüístico	Termos Lingüísticos	Substantivos – aluno, professor Adjetivo – inteligente. Verbo – cursar

FIGURA 13 – Níveis de representação para o conhecimento

Pode-se notar que o conhecimento representado pelas sentenças, em lógica de primeira ordem, podem ser expressos em linguagem natural como *todo aluno pertence ao corpo acadêmico e existem alunos que são inteligentes*. A tradução de sentenças de primeira ordem para a linguagem natural é facilitada pelo uso de palavras da linguagem natural como predicados (aluno, corpo acadêmico e inteligente). Este é um artifício muito usado pelos projetistas de bases de conhecimento, que esperam, desta forma, vincular a teoria formal com o significado pretendido. No entanto, como salienta Guarino (1995b) este objetivo não pode ser alcançado desta forma, pois as palavras em linguagem natural não possuem uma semântica precisa devido à ambigüidade e à imprecisão de seus elementos. Por exemplo, pode-se utilizar a palavra *livro* como um símbolo na linguagem de representação do conhecimento com o intuito de associar o símbolo ao conceito de livro. No entanto, a palavra livro é usada tendo mais de um sentido. O sentido associado à palavra na sentença *Pedro emprestou o livro* é diferente do sentido associado a palavra na sentença *Eu li o livro que Ana escreveu*. No primeiro caso, a palavra denota o conceito de livro como um objeto concreto, ou seja, no sentido de exemplar, identificado no tempo e no espaço. Já o segundo caso, a palavra denota o conceito de livro como um objeto abstrato, ou seja, no sentido de um produto intelectual, identificado por um autor e título.

No *nível epistemológico*, é capturada a estrutura do conhecimento por meio de primitivas de estruturação. Assim, no exemplo da FIG.13, *aluno* é declarado como subclasse do *corpo acadêmico* e *inteligente* como uma qualidade. Esta visão, provavelmente coincide com a interpretação intuitiva da maioria das pessoas, que vêem *inteligente* como um atributo de um objeto. No entanto, nada impede que outra pessoa tenha uma outra visão e interprete *inteligente* como um objeto e *aluno* como qualidade. No entanto, é preciso fixar a interpretação, para que seja possível produzir bases de conhecimentos que possam ser reutilizáveis e compartilháveis. Segundo Guarino (1995b) a fixação da interpretação depende do *nível ontológico*, onde são estabelecidos os compromissos ontológicos. Assim interpretações espúrias, como *aluno* ser uma qualidade de *aluno*, podem ser bloqueadas. No exemplo da FIG.13 *aluno* é definido como pertencendo à categoria *Objeto material*, enquanto que *inteligente* é definido como pertencendo à categoria *qualidade*. Estas categorias fazem parte das categorias propostas por Dahlberg (1978b). Os compromissos ontológicos podem ser registrados em qualquer linguagem suficientemente poderosa que permita a reificação⁸, como a Lógica de Primeira Ordem.

No *nível conceitual* as primitivas possuem uma interpretação cognitiva bem definida baseada em conceitos elementares independentes de linguagem, como ações ou papéis temáticos. No *nível lingüístico* as primitivas são associadas diretamente à categorias lingüísticas, como nomes e verbos.

Outro pesquisador que faz a distinção entre o aspecto ontológico e o epistemológico é Poli (Poli, 2001). Para ele, uma ontologia “é a teoria de objetos. E é assim para todo tipo de objeto, concreto e abstrato, existente e não-existente, real e ideal, independente e dependente. Qualquer que sejam os objetos que se esteja lidando, a ontologia é a teoria deles.” Partindo desta

⁸ O termo reificação vem do latim *res* ou coisa. Uma linguagem que permita reificação é uma linguagem que permita transformar predicados em termos. Por exemplo, podemos expressar em lógica de primeira ordem que *Pedro é homem* por meio da fórmula *homem(Pedro)* ou através da fórmula reificada *propriedade(Pedro, Homem)*. Note que o predicado *Homem* na primeira fórmula tornou-se um termo na segunda.

visão, Poli argumenta que existe um equívoco em grande parte das definições sobre ontologia, confundido-a com a *epistemologia* que é a teoria dos diferentes tipos de conhecimento e das formas que são usados. Segundo ele, a diferença entre esses dois enfoques se torna evidente quando são listados os conceitos da ontologia e os da epistemologia. Os conceitos ontológicos são: objetos, processo, particular, individual, todo, parte, evento, propriedade, qualidade, estado, etc. Já os conceitos epistemológicos são: crenças, conhecimento, conhecimento incerto, conhecimento errado, etc.

4.2.3.6 Linguagens para representação de ontologias

O registro dos comprometimentos ontológicos deve ser feito por meio de uma linguagem formal, com uma semântica bem definida. Segundo Guarino(1995b), qualquer linguagem formal com capacidade de reificação é adequada para o registro dos comprometimentos ontológicos. Devido a esses requisitos utilizou-se inicialmente a lógica de primeira ordem ou linguagens derivadas, como KIF (*Knowledge Interchange Format*) (Genesereth e Fikes, 1992) e lógica das descrições (Nardi e Brachman, 2002). No entanto, em razão da aridez da notação mais comumente utilizada na lógica matemática e da popularização de outras notações mais simples para a descrição de meta-dados, como XML e RDF, atualmente, tem-se optado com maior frequência pela utilização destas para o registro dos comprometimentos ontológicos.

Uma outra linguagem que tem obtido um grande número de adeptos para a representação de ontologias é a OIL (*Ontology Inference Layer*) (Fensel *et al.*, 2000). Seus autores a propõem como uma linguagem padrão para ontologias em aplicações baseadas na Web e foi adotada para servir de substrato para a implementação da Web semântica proposta por Tim Berners-Lee. Devido ao seu rápido crescimento como linguagem de representação de ontologia, decidiu-se utilizá-la para ilustrar como uma ontologia é apresentada na ciência da computação. Antes, no entanto, é

preciso descrever, informalmente, algumas características da linguagem OIL. OIL possui uma semântica baseada em lógica das descrições, primitivas de modelagem baseadas em orientação à objetos (classe, superclasse e atributos) e duas variações de sintaxe: uma baseada em XML e outra baseada em RDF. Em relação aos níveis de representação do conhecimento a linguagem OIL se ocupa de descrever dois níveis. O primeiro nível, denominado de *nível meta*, é onde as definições ontológicas são fornecidas. Segundo Fensel *et al.* (2000), é neste nível que a terminologia é definida, descrevendo um vocabulário estruturado com uma semântica bem definida. OIL foi projetada tendo em mente este nível. O segundo nível, denominado de *meta meta-nível*, está relacionado com a descrição de características tais como autor, nome, assunto, formato, etc. Para representação dos metadados utilizou-se do padrão DublinCore, versão 1.1. O exemplo a seguir ilustra a representação destes dois níveis por meio da linguagem OIL.

```

ontology-container
title "Animais africanos"
creator "Ian Horrocks"
subject "animal, alimentação, vegetarianos"
description "Um exemplo didático de uma ontologia descrevendo animais africanos"
description.release "1.01"
publisher "I. Horrocks"
type "ontology"
format "pseudo-xml"
format "pdf"
identifier "http://www.cs.vu.nl/~dieter/oil/TR/oil.pdf"
source "http://www.africa.com/nature/animals.html"
language "OIL"
language "en-uk"
relation.hasPart "http://www.ontosRus.com/animals/jungle.onto"
ontology-definitions
  slot-def come
    inverse é-comido-por
  slot-def tem-parte
    inverse é-parte-de
    properties transitivo
  class-def animal
  class-def planta
    subclass-of NOT animal
  class-def árvore
    subclass-of planta
  class-def ramo
    slot-constraint é-parte-de

```

```

has-value árvore
class-def folha
  slot-constraint é-parte-de
  has-value ramo
class-def defined carnívoro
  subclass-of animal
  slot-constraint come
  value-type animal
class-def defined herbívoro
  subclass-of animal
  slot-constraint come
  value-type
  planta OR
  (slot-constraint é-parte-de has-value planta)
class-def girafa
  subclass-of animal
  slot-constraint come
  value-type folha
class-def leão
  subclass-of animal
  slot-constraint come
  value-type herbívoro

```

FIGURA 14 – Exemplo de representação de ontologia por meio da linguagem OIL
 FONTE : FENSEL et al. 2000, p. 6.

Apenas a título de comparação informal sobre o poder de expressividade entre a linguagem OIL e a linguagem usada para elaboração de tesouros, é mostrado na FIG. 15 a conversão do exemplo da FIG. 14 para a forma de tesouros. É importante ressaltar que o exemplo da FIG. 15 não vale como um exemplo típico de tesouros, uma vez que não foi construído segundo a metodologia preconizada. Vale apenas para uma comparação preliminar sobre o poder de expressividade das duas linguagens no sentido de OIL para tesouros. O exemplo foi elaborado com o uso dos indicadores de relação propostos em Gomes (1990), e apenas a parte alfabética é mostrada. A relação “come” foi substituída no tesouros pelo termo “ALIMENTAÇÃO”

PARTE ALFABÉTICA

ALIMENTAÇÃO
 TA HERBÍVORO
 TA CARNÍVORO
 TA ANIMAL

ANIMAL
TE GIRAFA
TE LEÃO
TE HERBÍVORO
TE CARNÍVORO
TA PLANTA
TA ALIMENTAÇÃO
ÁRVORE
TG PLANTA
TEP RAMO
CARNÍVORO
TG ANIMAL
TA HERBÍVORO
TA ALIMENTAÇÃO
FOLHA
TGP RAMO
GIRAFA
TG ANIMAL
TA HERBÍVORO
HERBÍVORO
TG ANIMAL
TA PLANTA
TA ALIMENTAÇÃO
LEÃO
TG ANIMAL
TA CARNÍVORO
PLANTA
TE ÁRVORE
RAMO
TGP ÁRVORE
TEP FOLHA

FIGURA 15 – Tradução da ontologia para tesouro

Em uma análise informal pode-se observar que na linguagem OIL é mais fácil registrar certas propriedades das relações do que com o uso de tesouros. Por exemplo, não foi possível expressar, por meio de tesouros, que carnívoro é um animal que come animais. No entanto, tal detalhe de associação não é tão significativo quando o usuário está buscando um documento. Por outro lado, a relação parte-de não precisa ser declarada nos tesouros, uma vez que já faz parte do seu repertório de relações. Esta comparação inicial parece indicar que a linguagem OIL permite especificar

melhor as relações e, desta forma, melhor discriminá-las. É possível que algumas linguagens para tesauros incluam indicadores de relações que discriminem mais relações, porém, se não houver uma forma de discriminá-las conforme a necessidade, sempre existirão restrições que não poderão ser expressadas. Outra vantagem da descrição formal das relações na linguagem OIL é a possibilidade do uso de mecanismos automáticos de inferência para realização de deduções sobre o conhecimento, por meio das relações descritas.

4.2.3.7 Considerações sobre a ontologia na ciência da computação e inteligência artificial

Pode-se observar que, mesmo em uma única área, como a IA, existem divergências sobre o que é uma "ontologia". Existem divergências sobre o nível de atuação (conceitual ou simbólico), grau de formalismo (formal, semi-formal ou informal), abrangência (geral ou relacionada com um domínio específico) e quanto à natureza da entidade que é descrita (objeto ou conceito). Parte destas divergências podem ser creditadas à pressão pelo desenvolvimento de novas aplicações, que tem atropelado a necessidade do desenvolvimento de um embasamento teórico mais cuidadoso e a consolidação da terminologia para a área. A visão da ontologia como uma tecnologia, ainda está em suas primeiras etapas, mas existe uma imensa quantidade de estudos realizados no âmbito da filosofia e terminologia que poderiam ser levados em consideração na criação e denominação dessas estruturas.

Após esta análise pode-se afirmar que a ontologia na IA é um objeto distinto da ontologia na filosofia e com ela guarda poucos pontos em comum. Parte desta distinção pode ser atribuída à uma diferença de objetivo entre as duas ciências: enquanto a filosofia preocupa-se com o que existe no mundo, para a IA o que existe é o que pode ser representado em um computador (objetos virtuais). Teria sido melhor utilizar outro nome, para denominar este instrumento de representação do conhecimento, uma vez que a adição de

novos significados a uma palavra dificulta a comunicação entre as áreas. Talvez tivesse sido melhor usar outro termo, como por exemplo, *modelo conceitual*, apesar de seu uso ser contestado por Welty e Guarino (2001), uma vez que ontologias devem ser independentes de tarefas e modelos conceituais são orientados a tarefas. Outra possibilidade é o uso do termo *teoria ontológica*, introduzido por Bunge (1977, p.11), definido como “uma teoria que contém e inter-relaciona categorias ontológicas ou conceitos genéricos representando componentes ou características no mundo”.

No entanto, a palavra ontologia já está estabelecida dentro da IA e, após a publicação de inúmeros artigos e livros utilizando-a, não é possível sugerir outro termo. O máximo que se pode esperar é que se deixe claro, quando necessário, a distinção entre os dois significados, como é feito em (Valente, 1995), onde ontologia na IA é denominada de *AI ontology*.

5 METODOLOGIA

Esta é uma pesquisa que busca um esclarecimento terminológico e a caracterização de dois instrumentos utilizados na organização do conhecimento: tesouros e ontologias.

Todos os subsídios teóricos e materiais utilizados foram coletados de fontes bibliográficas relacionadas às temáticas estudadas, englobando as áreas de filosofia, ciência da informação, ciência da computação, nesta mais especificamente, a subárea da inteligência artificial.

Na ciência da informação como na ciência da computação têm-se tanto os subsídios teóricos como os subsídios materiais, as definições, que serão analisadas por meio de um método. Já em relação à filosofia acha-se presente neste texto apenas os subsídios teóricos, destinados ao fornecimento dos elementos para a análise qualitativa.

As definições sobre tesouros e ontologias serão estudadas aplicando-se o método analítico-sintético.

As técnicas adotadas são: análise quantitativa e qualitativa do conteúdo. O uso da análise qualitativa em conjunto com a quantitativa enriquece a análise dos dados que se pretende compreender. Os aspectos quantitativos são pontos de partida e de apoio para a análise e os aspectos qualitativos fornecem uma melhor compreensão dos dados obtidos.

A técnica aqui proposta também pode ser vista como uma técnica de *análise de conteúdo*. Análise de conteúdo pode ser entendida como "uma técnica de pesquisa para descrição objetiva, sistemática e quantitativa do conteúdo evidente da comunicação" (Berelson *apud* Marconi & Lakatos, 1982, p. 99). Já para Ander-Egg (1978, p.178) é "a técnica mais difundida para investigar o conteúdo das comunicações de massas, mediante a classificação, em categorias, dos elementos da comunicação". Assim, em relação a este trabalho o conteúdo das definições, tanto de ontologias quanto de tesouros, seriam analisados por meio de categorias sistemáticas,

emergentes das definições estudadas que levariam a resultados quantitativos.

Berelson apresenta uma série de princípios específicos em relação às questões que podem ser estudadas, ao aplicar esta técnica, tais como: questões a respeito das características do conteúdo; questões relativas ao produtor ou motivos de conteúdo; e questões em relação à audiência ou efeitos de conteúdo. Neste trabalho se justificaria o primeiro item: questões a respeito das características do conteúdo tornando possível comparar níveis dos dois conceitos: tesouros e ontologias.

Para tanto a análise de conteúdo deve seguir algumas restrições como menciona Selltiz *et al.* (1965, p. 392):

1. as categorias da análise usadas para classificar o conteúdo são definidas clara e explicitamente para que outros indivíduos possam aplicá-las ao mesmo conteúdo, a fim de verificar as conclusões;
2. o analista não é livre para selecionar e registrar simplesmente aquilo que chama sua atenção, por ser interessante, mas deve classificar metodicamente todos os assuntos importantes, em sua amostra; e
3. certo processo quantitativo é usado para proporcionar a média da importância e ênfase da matéria de várias idéias verificadas e para permitir confrontos com outras amostras do material.

Esta etapa seria a fase quantitativa compreendendo uma análise objetiva, que permitirá uma comparação dos resultados alcançados. A etapa seguinte seria uma discussão qualitativa dos resultados obtidos na fase quantitativa sobre os objetos estudados, utilizando-se da interpretação e discussão dos resultados alcançados tendo como subsídio o conhecimento prévio sobre o termo tal como expresso em capítulo específico.

5.1 Método Analítico-Sintético

O método proposto por Dahlberg (1978b) para análise e estruturação de conceitos baseia-se na análise das proposições verdadeiras sobre o conceito, gerando uma hierarquia de características, que tem como característica mais genérica uma categoria, e na síntese destas características na forma de um termo ou nome cujo significado está estabelecido precisamente por meio de uma definição. Por envolver tanto uma etapa de análise como uma etapa de síntese o método é denominado de analítico-sintético. Este método provê uma forma segura para compressão da intensão de um conceito para sua inserção em uma estrutura de conceitos de um dado domínio.

Para cumprir o objetivo geral de determinar o significado dos termos “ontologia” e “tesauro” dentro nas áreas descritas é preciso aplicar um método que permita o registro dos significados subjacentes a cada definição corrente e sua posterior comparação. Por esta razão, o método escolhido para essa tarefa é método analítico-sintético.

O método analítico-sintético tem sido empregado dentro da ciência da informação em situações que envolvem a construção de estruturas conceituais, como a construção de tesauros, como em Gomes (1990), ou em situações que envolvam a compressão e definição adequada de um conceito como em Alvarenga (1993).

Em seu trabalho Alvarenga utiliza-se do método analítico-sintético para analisar e propor uma definição para o conceito de “publicações oficiais”. Devido a analogia com o presente trabalho, os passos utilizados por Alvarenga no desenvolvimento de seu trabalho foram adaptados para esta pesquisa, incluindo-se a aplicação do KWIC para separar os termos constantes das definições.

A seguir apresenta-se, de forma detalhada, os passos da pesquisa apresentada neste trabalho.

- Levantamento em separado das definições e conceitos sobre ontologias e tesouros. Nesta etapa foram extraídas da literatura as definições e conceitos sobre ontologia na área da ciência da computação e de tesouros na ciência da informação.
- Submissão das definições a um programa para produção de uma ordenação alfabética de palavras de acordo com sistema de indexação KWIC. KWIC (*Keywords in context*) é uma forma de análise de texto onde as palavras são destacadas e listadas alfabeticamente, mantendo toda a frase e preservando o contexto da ocorrência da palavra, ou seja, as palavras que antecedem e sucedem a palavra destacada. Para essa fase foram eliminadas palavras não significativas (*stopwords*). No resultado do KWIC também aparece o número total de ocorrências da palavra.
- Agrupamento dos termos em categorias ou classes amplas. As categorias serão determinadas pelas semelhanças entre os atributos. Esta fase demanda às vezes uma tentativa de juntar num mesmo grupo atributos semelhantes, mesmo utilizando-se de termos diferentes, mas que dizem a mesma coisa.
- Discussão sobre as categorias. Nesta etapa discute-se as categorias escolhidas bem como os atributos e características a elas relacionadas.
- Transcrição em uma matriz de ocorrência dos conceitos ocorridos em cada categoria tanto sobre tesouro como sobre ontologia. Esta matriz será fundamental para fazer análise comparativa do conteúdo proposto.
- Análise das definições e conceitos segundo a ocorrência das categorias determinadas na fase anterior. Desta forma é possível verificar diferenças e semelhanças entre os dois conceitos.

- Uma discussão ampla dos resultados à luz do conhecimento prévio compreende a abordagem qualitativa prevista nessa pesquisa levando a conclusões.

5.2 Processo de seleção dos textos

Devido à grande quantidade de documentos não seria possível analisar toda a bibliografia relacionada com os temas em uma pesquisa de mestrado. Portanto, foi necessário o estabelecimento de critérios norteadores para a seleção dos documentos. Optou-se por selecionar os textos mais relevantes encontrados na literatura.

A atribuição de relevância a documento é uma tarefa difícil e envolve critérios polêmicos e/ou subjetivos. Acrescente-se a isso a necessidade de estabelecer critérios suficientemente sólidos para tornar o trabalho coerente e aceito pelas duas comunidades acadêmicas envolvidas.

Na fase de elaboração do projeto de pesquisa foram contatados alguns especialistas que, em consultas informais, indicaram os autores que mais se destacavam sobre o tema. Na área da ciência da computação contou-se com o auxílio de professores do Departamento de Informática da Universidade Federal de Viçosa. No caso dos tesouros, como trata-se de um instrumento tratado exaustivamente na ciência da informação, não foi necessário recorrer a especialistas de outras unidades acadêmicas, fora da Escola de Ciência da Informação da UFMG.

Pode-se observar que existe uma diferença na natureza do material colhido nas duas áreas. Na ciência da informação o termo tesouro possui um significado mais consolidado em razão dos anos de pesquisa e de utilização. Existem inúmeros documentos, artigos científicos, normas e manuais que tratam do assunto sob diversos ângulos. Já na ciência da computação pode-se observar que o termo ontologia encontra-se em fase de amadurecimento. O termo possui várias definições e ainda não existe uma metodologia amplamente aceita que guie a elaboração de ontologias. Logo, no caso da

pesquisa sobre ontologia, tornou-se óbvia uma maior necessidade do estabelecimento de critérios claros para a seleção dos documentos e definições sobre o assunto.

O critério adotado para a seleção de artigos da ciência da computação, relacionados com ontologias, para a fase de desenvolvimento da pesquisa, foi o número de citações recebidas. A frequência de citações foi obtida por meio de consultas ao site da CiteSeer (Lawrence *et al.*, 1999). CiteSeer (também conhecida por ResearchIndex) é uma biblioteca digital de literatura científica que tem por objetivo auxiliar na disseminação de literatura científica relacionada principalmente com a ciência da computação. A biblioteca é mantida pelo NEC Research Institute e pode ser acessada no endereço eletrônico *citeseer.org*. Dentre os serviços fornecidos pelo CiteSeer destacam-se:

- Indexação autônoma de citações – o CiteSeer analisa, por meio de um sistema computacional autônomo, documentos em formato eletrônico, retirando citações e armazenando-as em um índice para consultas posteriores. Os documentos podem ser fornecidos pelos próprios autores ou buscados automaticamente pelo sistema, na Internet. O sistema opera de forma completamente autônoma, ou seja, sem nenhuma intervenção humana. O sistema é capaz de analisar diversas formas de citação. Lawrence (Lawrence *et al.*, 1999)(Lawrence, 2001) apresenta maiores detalhes sobre o funcionamento do mecanismo de indexação autônoma de citações do CiteSeer.
- Análises estatísticas de citações para todos os documentos citados.
- Busca por palavras chaves com resultados ordenados por data, citações ou citações esperadas.

Em uma busca no CiteSeer utilizando a palavra “ontologies” foram retornados documentos que totalizaram 2674 citações. Os artigos foram

ordenados por uma função que leva em conta o número de citações e a data de publicação do artigo, sendo que artigos mais recentes recebem um peso maior. A FIG. 16 mostra um segmento do resultado da consulta contendo os três artigos melhor classificados. A listagem dos resultados continha no final o gráfico da FIG 17. que mostra a distribuição das citações em função do ano de publicação dos artigos.

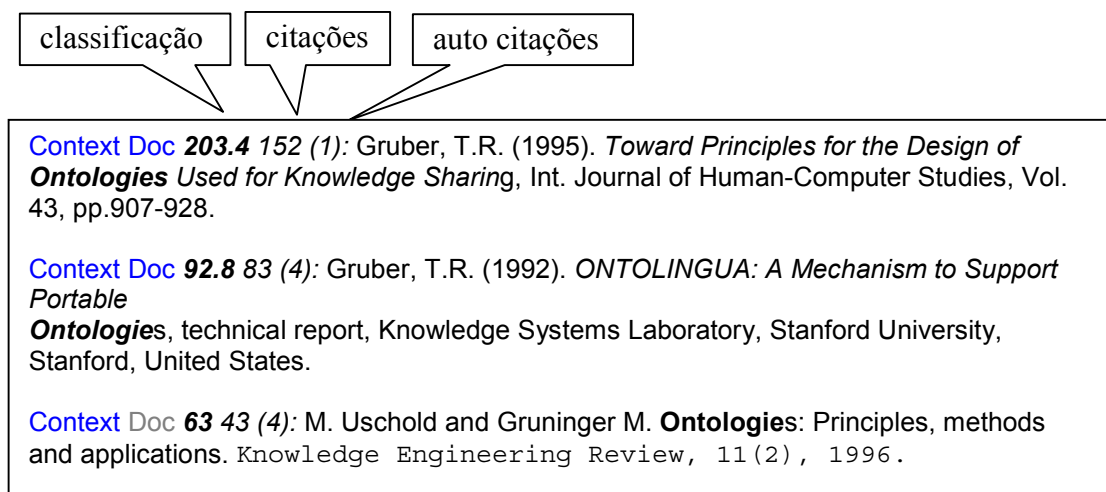


FIGURA 16 - Trecho do resultado da consulta usando a palavra “ontologies”

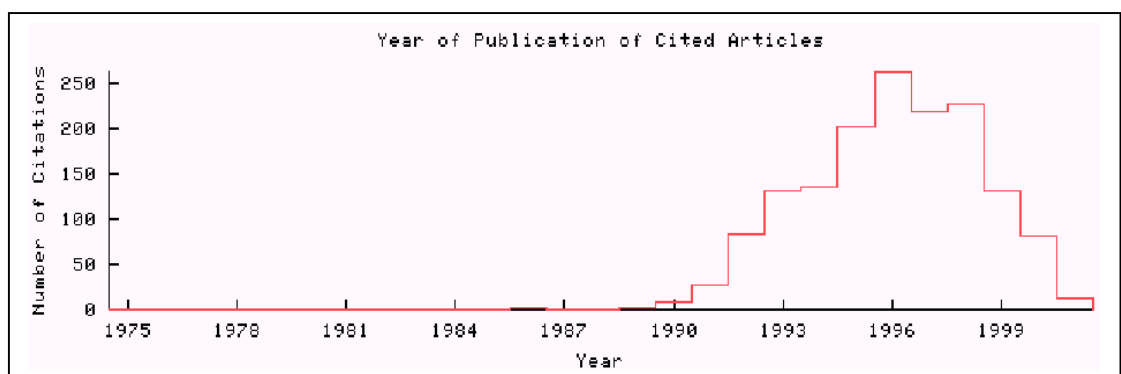


FIGURA 17 - Número de citações em função do ano de publicação dos artigos

A listagem retornada pela pesquisa realizada no CiteSeer revelou que após o artigo de posição 196 da listagem se seguiam artigos com apenas uma citação, demonstrando uma concentração de citações em poucos artigos. Além disso, alguns dos artigos mais citados pertencem a um mesmo autor, como pode ser observado na FIG. 16, o que parece indicar uma

concentração de citações em poucos autores. Nota-se também, um aumento do número de citações a partir da década de 90, mostrando que foi nesta época que surgiu o interesse pelo tema na ciência da computação. Pode-se observar também que, devido a erros de formatação, alguns artigos aparecem na listagem mais de uma vez, sendo interpretados pelo sistema como artigos diferentes. No entanto, este fato não teve qualquer influência sobre os resultados da pesquisa, exigindo apenas uma maior atenção na seleção dos artigos.

Em uma análise do conteúdo dos artigos pode-se notar que as pesquisas podiam ser classificadas, em sua maioria, em: pesquisas para desenvolvimento de ferramentas para criação de ontologias; metodologias para o desenvolvimento de ontologias; e aplicações baseadas em ontologias. Na primeira categoria enquadram-se produtos como Ontolingua (Gruber, 1992,1993), ODE (Blazquez et al., 1998) Tadzebao e WebOnto (Domingue, 1998) que podem ser usados para criar, manter e compartilhar ontologias. Na segunda categoria tem-se metodologias como a proposta por Guarino (1997), Gruninger e Fox (1995) e Uschold (1996). Na terceira categoria tem-se o uso de ontologias na gerência de conhecimento (Benjamins et al., 1998), no desenvolvimento de sistemas baseados em conhecimento (Van Heijst et al., 1997) e na integração de empresas (Fox e Gruninger, 1994). Os textos foram separados segundo esta classificação e selecionados aqueles que propunham definições originais para o termo ontologia.

Nota-se que o termo “ontologia” na área da filosofia não foi analisado por ser um termo bem consolidado pela comunidade de filósofos. Porém o entendimento do termo na filosofia foi utilizado nas discussões sobre as definições do termo na ciência da computação, visando verificar se existe uma preocupação, por parte da comunidade da ciência da computação, de fundamentar o termo filosoficamente.

5.3 Material

Neste trabalho o material empírico são as definições contidas nos textos selecionados sobre tesouros na CI e ontologias na CC. A análise das definições utilizadas neste estudo podem ser vistas como uma tentativa dos pesquisadores de estabelecer os elementos ou características, entendidos por eles como essenciais para a compreensão do objeto, sintetizando-os em um enunciado.

O processo de extração foi o seguinte: no caso do termo “ontologia” as definições foram retiradas dos textos mais citados segundo busca realizada no CiteSeer. Já para o termo “tesauro” as definições foram retiradas de normas e de textos sobre o tema produzidos por autores respeitados como Lancaster (1986 e 1987) e Gomes (1990). No caso da ontologia a maior dificuldade se deveu ao fato de que todos os textos são em inglês e, portanto, houve a necessidade de realizar uma tradução nas definições. No caso dos tesouros esta dificuldade não existiu, devido a existência de autores consagrados que escrevem em português. Além disso, existem publicações em português de normas com definições sobre tesauro e traduções de textos de autores clássicos.

Para o termo "tesauro" foram encontradas seis definições (ANEXO II), sendo que duas definições constam do mesmo documento da UNESCO. A UNESCO apresenta a definição no sentido estrutural e no sentido funcional, e neste trabalho, ambas foram utilizadas. A quantidade de definições encontradas para "tesauro" foi a metade da quantidade de definições encontradas para o termo ontologia na ciência da computação. Provavelmente, isto se deve ao fato do termo tesauro estar mais consolidado na ciência da informação, fazendo com que exista uma menor quantidade de variações.

Para o termo "ontologia" foram utilizadas treze definições (ANEXO VII) extraídas de vinte e nove artigos retornados da pesquisa realizada no CiteSeer. A pesquisa foi realizada com a palavra “*ontologies*” e os artigos foram ordenados por frequência de citação. A razão de terem sido extraídas

apenas treze definições de vinte e nove artigos deve-se ao fato de que nem todos os artigos apresentam definições, sendo que os que não apresentam definições adotam a proposta por Gruber. Esta é uma evidência da aceitação que a definição proposta por Gruber está obtendo dentro da ciência da computação. Além disso, alguns dos artigos são de um mesmo autor que, portanto, usa uma mesma definição em todos artigos.

As definições selecionadas foram submetidas ao programa ptx (*permuted indexer*), versão beta 0.3, para produção do KWIC. O ptx é disponibilizado livremente sob a licença da Fundação de Software Livre, *Inc.* e distribuído junto com o sistema operacional Linux. As listagens obtidas por meio do ptx foram trabalhadas, gerando duas novas listas contendo conjuntos de termos agrupados em categorias para uma posterior análise quantitativa e qualitativa.

5.4 O KWIC

KWIC (*KeyWord In Context*) é uma ferramenta de indexação, enquadrada dentro da classe de ferramentas denominadas de *concordância*. Nas palavras de Gilchrist (1971), uma concordância é "um índice para palavras ou frases de um livro. ... A palavra é, usualmente, exibida com o seu contexto". Sardinha (2000) reforça a necessidade da apresentação do contexto de ocorrência da palavra quando diz que uma concordância "consiste numa listagem dos co-textos (palavras ao redor) nos quais um dado item (palavra isolada, composta, estrutura, etc.) ocorre." Segundo Sardinha, os tipos mais comuns de concordância são o KWIC e o KWOC (*KeyWord Out of Context*). No caso da concordância KWIC cada palavra relevante é mostrada no centro da listagem, ordenada alfabeticamente, ladeada pelas palavras antes e depois dela no texto. Já a concordância KWOC é uma listagem de títulos ordenados pelas palavras relevantes que neles ocorrem. O exemplo abaixo mostra o uso KWIC no destaque das palavras dentro de seu contexto.

Ontologia como um sistema conceitual relacionado a uma	base	de conhecimento
Tesouro é uma	linguagem	especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.

FIGURA 18 - Exemplo do uso KWIC

Os instrumentos de concordância não são recentes. Gilchrist (1971) menciona o uso de concordância feito por Cruden em 1738 sobre o texto da Bíblia. Ele menciona também que em 1856, Crestadoro sugeriu o uso de KWIC sobre títulos em um trabalho denominado de *The Art of Making Catalogues*. Atualmente, os sistemas de concordância, principalmente o KWIC, são utilizados para análise de texto, em estudos de lexicometria e no auxílio a buscas.

A concordância KWIC foi a selecionada para ser usada no presente estudo devido a sua facilidade para evidenciar as palavras dentro de seu contexto (neste caso as definições). O uso desta ferramenta, permitiu uma visualização mais rápida das palavras que ocorrem nas definições.

5.5 Escolha dos termos

A seleção de termos obedeceu a abordagem onomasiológica⁹, ou seja, selecionou-se os termos que denotassem conceitos do domínio. Como consequência desta abordagem, alguns termos compostos não foram desmembrados em seus termos constituintes, uma vez que denotam conceitos relevantes no domínio ou unidades lexicais indissociáveis. Por outro lado, outros termos compostos foram desmembrados em seus termos

⁹ O estudo das expressões de que dispõe uma língua para traduzir determinada noção, e que parte, pois, do significado para estudar o significante. (Ferreira, 1975)

constituintes uma vez que estes denotavam conceitos existentes no contexto. Por exemplo, o termo “base de conhecimento” não foi desmembrado em “base” e “conhecimento”, uma vez que o termo composto denota uma entidade conhecida pela comunidade de inteligência artificial. Já o termo “linguagem de informação” foi desmembrado nos termos “linguagem” e “informação” uma vez que o termo composto não representa algo mais significativo do que seus termos constituintes.

Utilizou-se definições de dicionário para descrever parte dos termos privilegiados da rotação pelo sistema KWIC. Os termos restantes foram descritos por meio de definições de autores da área ou por meio de descrições que capturam o significado do termo no contexto. Esta ação de ir buscar a noção do objeto a ser categorizado é compreendida como uma abordagem pragmática onde se registra interpretações adequadas ao contrário da abordagem epistemológica, onde os conceitos são definidos no nível abstrato. No manual de tesouros monolíngues é indicado que apenas os termos básicos sejam definidos por causa do despendimento de tempo nesta atividade, mas aqui tentou-se caracterizar todos os elementos utilizados.

Uma vez que o programa para geração do KWIC gera uma saída para palavras e não para termos, foi necessário uma edição da saída para incluir os termos que formassem uma unidade semântica. Isto foi feito, tomando-se cada palavra emitida pelo programa e examinando se denotava um conceito do domínio. Caso contrário, era feita uma composição com as palavras ocorrendo em sua vizinhança e uma nova verificação, em relação ao sentido, era realizada.

6 Resultados

6.1 Categorias

Após a seleção das definições e extração dos termos, estes identificados pelas unidades semânticas presentes, teve início a etapa de elaboração das categorias. Este é um processo cujo objetivo é juntar sob um mesmo grupo conceitos semelhantes representados por termos distintos. As categorias foram determinadas por meio das semelhanças entre as características inerentes aos termos em um processo de indução. Há de se considerar que as definições foram retiradas de áreas distintas, com terminologias próprias e, portanto, em determinadas situações, apesar de usarem termos distintos podem estar se referenciando a um mesmo conceito. O uso de categorias visa equalizar os termos, uma vez que os termos abrigados sob a mesma categoria compartilham do significado apontado por ela.

As categorias aplicadas no contexto de ontologias foram também aplicadas aos tesouros, sendo a categoria "agente" a única exceção, uma vez que não ocorreu no contexto das ontologias. Lista-se a seguir as categorias obtidas nesta etapa:

Objeto – agrupa os termos relacionados com os objetos concretos mencionados nas definições, tais como “bases de conhecimento” e “documentos”.

Espaço de conhecimento – reúne os termos relacionados com a delimitação do espaço de conhecimento, tais como “assunto” e “domínio”.

Conteúdo semântico – agrupa os termos associados a elementos de significado, tais como “conceitos” e “definições”.

Linguagem – reúne os termos relacionados com o tipo de linguagem, é usada pelos objetos que estão sendo definidos, tais como “linguagem natural”, “vocabulário controlado” e “vocabulário”. Os termos “linguagem especializada” e “vocabulário controlado” não devem ser desmembrados pois possuem uma conotação inequívoca, consagrada no contexto da CI.

Processos – agrupa os termos relacionados com os processos mencionados nas definições, tais como “indexação”, “descrição” e “restrição”.

Atributo – reúne os termos associados a propriedades dos objetos que estão sendo definidos, tais como “normalizada”, “formal” e “explícita”.

Sistematização – agrupa os termos relacionados com a organização e estruturação, tais como “sistema”, “esquema” e “estrutura”.

Agentes - agrupa os termos relacionados a agentes, tais como usuários.

6.2 Atribuição dos termos às categorias

A atribuição dos termos às categorias ocorreu simultaneamente à criação das categorias como parte do processo. No entanto, por questão de organização, esta associação é apresentada nesta seção. Como já foi mencionado na metodologia, a seleção de termos obedeceu a abordagem onomasiológica, ou seja, selecionou-se os termos que denotassem conceitos do domínio. Segue a relação dos termos agrupados por categorias, juntamente com suas definições ou com uma pequena descrição. As descrições foram utilizadas quando não foi encontrada uma definição que

capturasse o sentido dentro da área de estudo. As definições foram obtidas a partir de dicionários ou a partir de textos de autores da área.

6.3 Termos selecionados das definições sobre ontologia classificados por categorias

Categoria "Objeto"

Base de conhecimento – conjunto de representações de fatos sobre o mundo (Russel e Norvig, 1995). Um conjunto de proposições sobre uma determinada situação. É o produto da engenharia de conhecimento e, como tal, é um objeto material. Ocorre em duas definições: definições 4 e 10.

Entidades – aquilo que constitui a natureza fundamental ou a essência de uma coisa, aquilo que existe ou imaginamos que existe; ente, ser, Individualidade (dicionário Michaelis). Nas ontologias denotam os objetos, de qualquer natureza (concretos ou abstratos), que ocorrem no domínio. Ocorre na definição número 7.

Categoria "Espaço de conhecimento"

Assunto – aquilo de que se trata, que é objeto ou matéria de observação, consideração, atenção, interesse (Ferreira, 1975). Ocorre na definição número 1.

Domínio – possui o mesmo significado que assunto. Ocorre em sete definições: definições 1, 2, 3, 8, 9, 10 e 11 .

Categoria "Conteúdo semântico"

Conceitos – representação dum objeto pelo pensamento, por meio de suas características gerais (Ferreira, 1975). Assim, na computação, o conceito ocorre no nível do pensamento e atribui significado a um termo ou símbolo e,

por isso, deve ser enquadrado na categoria de conteúdo semântico. Ocorre na definição número 11.

Conceitualização (ões) – uma estrutura semântica que codifica as regras implícitas que restringem um segmento da realidade (Guarino e Giaretta, 1995). Ocorre em duas definições: 6 e 12.

Conhecimento – idéia, noção; informação, notícia (dicionário Michaelis). Ocorre na definição 2.

Definições – equivalência estabelecida entre um termo (v. definiendum) e outros termos conhecidos (v. definiens), e pela qual se cria um objeto determinado de pensamento (Ferreira, 1975). Ocorre em duas definições: 8 e 11.

Significados – sentido de uma palavra ou frase, o que uma coisa denota, sentido em que se emprega um termo; acepção (dicionário Michaelis). Ocorre na definição número 11.

Categoria "Linguagem"

Linguagem natural – qualquer linguagem de uso geral, escrita e/ou falada por uma comunidade humana (Ferreira, 1975). Ocorre na definição número 8, como uma ferramenta para expressar definições.

Lógica de descrições – subconjunto da lógica de primeira ordem “usada para descrever o conhecimento em termos de conceitos e restrições de relações que podem automaticamente derivar taxionomias de classificação” (Fensel *et al.*, 2001). Ocorre na definição número 9.

Termos – vocábulo ou locução que denomina conceito, prévia e rigorosamente definido, peculiar a uma ciência, arte, profissão, ofício (Ferreira, 1975). Unidade de comunicação que representa o conceito. Ocorre em seis definições: definições 3, 8, 9, 10, 11 e 13.

Vocabulário – o conjunto das palavras especializadas em qualquer campo de conhecimento ou atividade (Ferreira, 1975). Ocorre nas definições 1, 2, 11 e 13.

Categoria "Processos"

Descrição / descrever / descrevendo – ato ou efeito de descrever (Ferreira, 1975). Ocorre nas definições número 2, 7 e 10.

Especificação – ato ou efeito de especificar. Descrição rigorosa e minuciosa das características que um material, uma obra ou um serviço deverão apresentar. (Ferreira, 1975). Sua ocorrência nas definições mostra a existência de uma preocupação com o rigor da descrição . Ocorre em três definições: 6, 11 e 13.

Interpretações – ato ou efeito de interpretar. Explicação, comentário. Representação. (Ferreira, 1975). Ocorre na definição número 11.

Representação – ato ou efeito de representar(-se). Exposição escrita de motivos, etc. Coisa que se representa. Ato de representar o conteúdo concreto apreendido pelos sentidos, pela imaginação, pela memória ou pelo pensamento (Ferreira, 1975). Ocorre em três definições: 1, 2 e 5.

Categoria "Atributo"

Formal / formais – preciso, próprio, genuíno (Ferreira, 1975). Expresso em uma linguagem cuja semântica é precisa. Ocorre em três definições: 6, 7 e 8.

Explícita/ expressos – expresso formalmente; claro, desenvolvido, explicado (Ferreira, 1975). Ocorre nas definições número 6 (explícita) e na definição 9 (expressos).

Hierarquicamente – na orientação a objetos, organização de classes que indica suas subordinações com relação à herança (Ferreira, 1975). Ocorre na definição número 10.

Compartilhadas – ter ou tomar parte; participar; compartilhar (Ferreira, 1975). No contexto refere-se à propriedade de um conhecimento ser compartilhado por uma comunidade. Ocorre na definição número 12.

Declarativo – tipo de conhecimento onde “o conhecimento é especificado, mas o uso que será feito dele não é fornecido” (Rich, E., Knight, K., 1993). Ocorre na definição número 3.

Especializado – estudo especial de certa arte ou ciência (dicionário Michaelis). Ocorre na definição 1.

Propriedade – qualidade de próprio, qualidade especial; particularidade (Ferreira, 1975). Ocorre na definição 7.

Comportamentos – maneira de se comportar; procedimento, conduta (Ferreira, 1975). Ocorre na definição 7.

Restrições / restringe – Dar sentido restrito a (palavra ou frase). Conter dentro de certos limites, limitar(-se) (dicionário Michaelis). Nas ontologias há necessidade de restrições. Ocorre em três definições: 7, 8 e 11.

Categoria "Sistematização"

Esquema – conjunto de relações que, para fins operatórios (sic), é suficiente para caracterizar um sistema, sem, contudo, esgotar-lhe a complexidade. (Ferreira, 1975). Ocorre na definição número 10.

Estrutura /estruturados – a disposição dos elementos ou partes de um todo; a forma como esses elementos ou partes se relacionam entre si, e que determina a natureza, as características ou a função ou funcionamento do todo (Ferreira, 1975). Ocorre em duas definições: 10 e 11.

Sistema Conceitual – descrição não formal dos relacionamentos entre conceitos. Ocorre em duas definições: 4 e 5.

Modelo – representação simplificada e abstrata de fenômeno ou situação concreta, e que serve de referência para a observação, estudo ou análise, baseia-se numa descrição formal de objetos, relações e processos, e que permite, variando parâmetros, simular os efeitos de mudanças de fenômeno que representa (Ferreira, 1975). Ocorre na definição número 3.

Relações – operação que determina a agregação ou a conexão de dois objetos (Ferreira, 1975). Ocorre em cinco definições: 3, 7, 8 e 13.

Inter-relacionada – Relação mútua (Ferreira, 1975). Ocorre na definição número 11.

6.4 Termos selecionados das definições sobre tesouros classificados por categorias

Categoria "Objeto"

Documentos / Documentários – qualquer base de conhecimento, fixada materialmente e disposta de maneira que se possa utilizar para consulta, estudo, prova, etc. (Ferreira, 1975). Já documentário entendido como "Relativo a documentos, Que tem o valor de documento". O termo "Documentos" ocorre na definição 6 e "documentários" ocorre na definição 2

Informação – "... conjunto de dados, fatos ou símbolos com potencial para sentido" (Fabiosoff citada por Paim). Ocorre na definição 6.

Categoria "Espaço de conhecimento"

Domínio – âmbito de uma arte ou ciência (Ferreira, 1975). Na CI esta expressão é usada quando se faz um recorte do conhecimento, por exemplo, ao se construir tesouros, onde se estabelece um âmbito específico do conhecimento a ser analisado. Ocorre na definição 5.

Categoria "Conteúdo semântico"

Conceito – na CI, segundo o entendimento da teoria do conceito de Dahlberg (1978a) o conceito é visto como uma unidade de conhecimento, e portanto deve ser enquadrado na categoria de conteúdo semântico. Ocorre na definição número 3.

Conhecimento – idéia, noção; informação, notícia (dicionário Michaelis). Ocorre na definição 5.

Semanticamente/Semântica – estudo da relação de significação nos signos e da representação do sentido dos enunciados (Ferreira, 1975). Nocetti e Figueiredo (1978) afirmam que os vocabulários documentários normalizados vão se apresentar em quatro tipos de relações semânticas (relação de significação), que são: equivalência, hierarquização, associação e definição. Aqui entende-se semântica no sentido de significado ou da idéia do termo, em que este permitirá gerar relacionamentos entre termos. Ocorre em duas definições: 2 e 5.

Categoria "Linguagem"

Lista Autorizada – é entendido como uma lista controlada de termos (Lancaster, 1986). Ocorre em duas definições: 3 e 4.

Elementos Lingüísticos – em uma citação de Currás (1995) tem-se a seguinte menção: "um sistema lingüístico no qual os componentes principais são os termos". Portanto, elementos lingüísticos podem ser entendidos como os termos existentes numa linguagem documentária. Ocorre na definição 2.

Indicadores de Relação – Gomes (1990) menciona que os indicadores de relação são códigos que precedem termos para indicar as relações dos termos. Ferreira (1975) define código como cada uma das seqüências de signos (números, letras, etc.) usados para identificação e controle dos elementos de determinado conjunto ou âmbito, segundo um método ou sistema geral de classificação. Indica que se faz referência a uma língua como meio de comunicação entre indivíduos, e não como um conhecimento parcialmente herdado e parcialmente adquirido. Ocorre na definição 1.

Linguagem – faculdade de expressão audível e articulada do homem. Conjunto de sinais falados, escritos ou gesticulados de que se serve o homem para exprimir suas idéias e sentimentos. Qualquer meio que sirva para exprimir sensações ou idéias (dicionário Michaelis). Ocorre na definição 6.

Linguagem Especializada – Currás (1995) define como "conjunto de termos relativos a um assunto, que formavam parte de um vocabulário especializado, compunha uma linguagem, por sua vez, especializada. Ocorre na definição 2.

Linguagem Natural – Qualquer linguagem de uso geral, escrita e/ou falada por uma comunidade humana (Ferreira, 1975). Como parte da tarefa de organização dos documentos na CI, traduz-se os documentos em linguagem natural para uma linguagem documentária. Ocorre em uma definição: definição 6.

Linguagem de documentação – Currás (1995) define como "linguagem controlada - normalizada - usada com fins classificatórios, no sentido amplo desta expressão". Ocorre na definição 6.

Linguagem pós-coordenada – Currás (1995) define como "linguagem documentárias nas quais os termos que as compõem se coordenam em processo posterior à sua determinação, por exemplo, no momento de seu estabelecimento ou de seu uso". Ocorre na definição 2.

Léxicas / Léxicos – o vocabulário de uma língua (Ferreira, 1975). Ocorre na definição 4.

Termos / Termos (Simples ou Compostos) – (gramática) Vocábulo, dicção, palavra, expressão. (Lógica) Conceito representado por sua expressão verbal. (Dic. Michaelis). Na CI pode ser entendido a expressão verbal que tem uma relação unívoca com o conceito. Ocorre em três definições: 1, 2, 5.

Sintática – parte da semiologia que se interessa especificamente pelas relações entre os signos (Ferreira, 1975). Nocetti e Figueiredo (1978) afirmam que as relações ou funções sintáticas das linguagens documentárias são úteis aos documentalistas quando precisam localizar o nome documentário de um conceito. Este tipo de relacionamento evita surgir problemas de combinações falsas, também chamadas de malformações, no processo de recuperação. Portanto, entende-se sintática no sentido de que os signos utilizados na linguagem, que irão se relacionar,

devam ser utilizados de forma adequada para que se reduzam falsas combinações na recuperação de informação. Ocorre na definição 2.

Terminológico – relativo a terminologia, conjunto dos termos próprios duma arte ou duma ciência; nomenclatura (Ferreira, 1975). Ocorre na definição 6.

Vocabulário Controlado – Lancaster (1986) apresenta vocabulário controlado como um conjunto limitado de termos que devem ser usado para representar o assunto dos documentos. Tal vocabulário pode ser uma lista de cabeçalhos de assunto, um esquema de classificação, um tesouro, ou simplesmente uma lista de palavras-chaves aprovadas ou frases. Também em uma tradução e adaptação do texto acima, conceitua-se vocabulário controlado como uma lista controlada de termos ou como uma lista autorizada (linguagem de indexação, vocabulário padrão). É com este entendimento que se categorizou então lista autorizada. Ocorre nas definições 1 e 5.

Categoria "Processos"

Indexação — processo que consiste na indicação do conteúdo informativo de um documento através da determinação de um ou mais termos que representarão esse conteúdo. Ocorre na definição 4.

Tradução – processo de converter uma linguagem em outra (Ferreira, 1975). Para CI é o processo de traduzir as palavras da linguagem do documento para os termos indicados na linguagem do sistema. Ocorre na definição 6.

Controle – ato, efeito ou poder de controlar; domínio, governo (Ferreira 1975). Na CI busca-se controlar e restringir o significado dos vocabulários contidos nas listas. Ocorre na definição 6.

Categoria "Atributo"

Associativa – na CI é um tipo de relação não hierárquica onde um termo se relaciona a outro termo por associações que dependem do contexto considerado. Ocorre na definição de número 1.

Coordenada – na CI é entendido como a combinação de classes ou termos. Ocorre na definição 4.

Dinâmico – é a qualidade do poder de atualização dos termos. Ocorre na definição 5.

Equivalência – na CI são considerados equivalentes ou de significado quase idêntico quando os termos equivalentes correspondem a um mesmo conceito, sendo reunidos em categorias de equivalência. Ocorre na definição 1.

Específico – alguma coisa adaptada particularmente para o seu fim (Dic. Michaelis). Ocorre na definição 5.

Genericamente – No sentido de indicar relações de generalização. Ocorre na definição 5.

Heurísticas – procedimento pedagógico pelo qual se leva o aluno a descobrir por si mesmo a verdade que lhe querem inculcar (Ferreira, 1975). Na CI os sistemas de classificação, em consequência de sua organização, levariam os usuários a descobrirem as informações desejadas. Ocorre na definição 3.

Hierárquicas – As relações hierárquicas podem ser entendidas como de gênero/espécie e de parte/todo. As relações hierárquicas exprimem as relações de subordinação que são analisadas sob 4 aspectos: do geral para o específico (TG - termo genérico); do específico para o geral (TE - termo específico); do todo para a parte (TGP - termo genérico partitivo); e da parte para o todo (TEP - termo específico partitivo). Ocorre na definição 1.

Intuitivas – que se percebe por intuição; claro, manifesto, evidente, imediato (Ferreira, 1975). Na CI intuitiva ocorre no mesmo sentido que relações heurísticas. Ocorre na definição 3.

Normalizada – Currás (1995) define como "... controlada, a qual se aplicam normas logicamente prefixadas ...". Ocorre na definição 2.

Organizado – ordenado, metódico (Ferreira, 1975). Na CI os vocabulários são organizados em listas. Ocorre na definição 1.

Restrita – que se mantém dentro de certos limites; limitado (Ferreira, 1975). Ocorre na definição 6.

Categoria "Sistematização"

Ordem – boa disposição das coisas, cada uma no lugar que lhe corresponde; disposição das coisas cujo arranjo se subordina a um princípio útil, agradável ou harmônios (Dic. Michaelis). Ocorre na definição número 1.

Sistema – conjunto de princípios, coordenados entre si de maneira a formar um todo científico ou um corpo de doutrina (Dic. Michaelis). Ocorre na definição 6.

Relacionados / relação / relações – de que se faz relação, operação que determina a agregação ou a conexão de dois objetos (Ferreira, 1975). Ocorre em quatro definições: relacionados 2 e 5, relações 1 e 3.

Categoria "Agentes"

Usuário / usuários – aquele que, por direito proveniente de uso, frui as utilidades da coisa (Ferreira, 1975). Na CI o usuário possui um papel fundamental na tarefas da organização documental. Toda tarefa realizada nos centros de documentação ou bibliotecas tem por finalidade facilitar o usuário em sua busca por informações. Estes termos ocorrem nas definições números 3 e 6.

Indexadores - Na CI os indexadores são aqueles que indexam os documentos, ou seja, o indexador indica o conteúdo informativo de um documento através da determinação de um ou mas termos que representarão esse conteúdo. Ocorre na definição número 6.

6.5 Quantificação dos termos

A tabela a seguir mostra a ocorrência de cada termo nas categorias e se são originários das definições sobre ontologias e/ou tesouros. Desta forma é possível visualizar as sobreposições entre os conceitos tratados nas definições.

CATEGORIAS /TERMOS	ONTOLOGIA	TESAURO
CATEGORIA "OBJETO"		
BASE DE CONHECIMENTO	X	
DOCUMENTOS / DOCUMENTÁRIOS		X
ENTIDADES	X	
INFORMAÇÃO		X
CATEGORIA "LINGUAGEM"		
ELEMENTOS LINGÜÍSTICOS		X
INDICADORES DE RELAÇÃO		X
LÉXICAS (OS)		X
LINGUAGEM		X
LINGUAGEM DE DOCUMENTAÇÃO		X
LINGUAGEM ESPECIALIZADA		X
LINGUAGEM NATURAL	X	X
LINGUAGEM PÓS-COORDENADA		X
LISTA AUTORIZADA		X
LÓGICA DE DESCRIÇÕES	X	
SINTÁTICA		X
TERMINOLÓGICO		X
TERMOS	X	X
TERMOS COMPOSTOS		X
TERMOS SIMPLES		X
VOCABULÁRIO	X	
VOCABULÁRIO CONTROLADO		X
CATEGORIA "ESPAÇO DE CONHECIMENTO"		
ASSUNTO	X	
DOMÍNIO	X	X
CATEGORIA "CONTEÚDO SEMÂNTICO"		
CONCEITO (S)	X	X
CONCEITUALIZAÇÕES	X	
CONHECIMENTO	X	X
DEFINIÇÕES	X	
SEMÂNTICA / SEMANTICAMENTE		X
SIGNIFICADOS	X	
CATEGORIA "PROCESSOS"		

CONTROLE		X
DESCRIÇÃO / DESCREVER / DESCREVENDO	X	
ESPECIFICAÇÃO	X	
INDEXAÇÃO		X
INTERPRETAÇÕES	X	
REPRESENTAÇÃO	X	
TRADUÇÃO		X
CATEGORIA ATRIBUTO		
ASSOCIATIVAS		X
COMPARTILHADAS	X	
COMPORTAMENTOS	X	
COORDENADA		X
DECLARATIVO	X	
DINÂMICO		X
EQUIVALÊNCIA		X
ESPECÍFICO		X
EXPLÍCITA /EXPRESSOS	X	
ESPECIALIZADO	X	
FORMAL / FORMAIS	X	
GENERICAMENTE		X
HEURÍSTICAS		X
HIERARQUICAS/ HIERARQUICAMENTE	X	X
INTUITIVAS		X
NORMALIZADA		X
ORGANIZADO		X
PROPRIEDADE	X	
RESTRIÇÕES/ RESTRINGE/ RESTRITA	X	X
CATEGORIA "SISTEMATIZAÇÃO"		
ESQUEMA	X	
ESTRUTURA/ ESTRUTURADOS	X	
INTER-RELACIONADA	X	
MODELO	X	
ORDEM		X
RELAÇÃO / RELAÇÕES /RELACIONADOS	X	X
SISTEMA		X
SISTEMA CONCEITUAL	X	
CATEGORIA "AGENTES"		
USUÁRIO / USUÁRIOS		X
INDEXADORES		X

TABELA 1 - Distribuição dos termos por categoria

Pode-se observar que na categoria "objeto" não houve nenhum termo que fosse oriundo tanto das definições de tesouros quanto de ontologias. Já na categoria "linguagem" os termos "linguagem natural" e "termos" são oriundos dos dois contextos. Alguns termos são próximos, como é o caso de "vocabulário" do contexto de ontologias e "vocabulário controlado", no caso

de tesouros. Na categoria "espaço de conhecimento" o termo "domínio" é usado em ambos os contextos. Na categoria "conteúdo semântico" os termos "conceito" e "conhecimento" são usados em ambos os contextos e os termos "significados" (ontologia) e "semântica" (tesauro) possuem o mesmo sentido. Percebe-se, também, que na categoria "processos", não ocorreu coincidência na origem dos termos. Isto provavelmente se deveu à diferença de atividades entre as duas áreas. Apesar da categoria "atributo" abrigar um bom número de termos, apenas os termos "hierárquicas/ hierarquicamente" e "restrições/restringe/restrita" ocorreram tanto nas definições de tesauro com nas de ontologia. Na categoria "sistematização" o termo "relação/relações/relacionados" é usado em ambos os contextos. A categoria "agentes" não registrou ocorrência de termo oriundo do contexto de ontologias.

A tabela a seguir apresenta uma quantificação da ocorrência dos termos por categoria. A primeira coluna da tabela lista todas as categorias levantadas. A segunda e a terceira colunas registram o número de termos, retirados das definições de ontologia que ocorrem na categoria em números absolutos e em percentual respectivamente. A quarta e a quinta coluna registram o número de termos, retirados das definições de tesauro que ocorrem na categoria em números absolutos e em percentual respectivamente. Todas as ocorrências de termos repetidos em definições distintas são contabilizadas.

Categorias	Ontologia		Tesouro	
objeto	3	5%	3	6%
linguagem	12	18%	17	35%
espaço de conhecimento	8	12%	1	2%
conteúdo semântico	7	11%	4	8%
processo	10	15%	3	6%
atributo	14	22%	12	25%
sistematização	11	17%	6	12%
agentes	0	0%	3	6%

TABELA 2 - Frequência de Ocorrência dos termos por categoria.

Nota-se, no caso das definições de tesouro, uma concentração de termos nas categorias linguagem e atributo, com uma frequência de ocorrência de termos bem maior, quando comparada com a registrada em outras categorias. Já em relação às definições sobre ontologia nota-se uma melhor distribuição, com uma ênfase maior nas categorias atributo, sistematização e linguagem. A categoria agentes não abrigou nenhum termo oriundo das definições sobre ontologia. Nota-se também que a categoria processo registrou uma frequência de ocorrência de termos oriundos das definições sobre ontologia bem maior do que a frequência de ocorrência de termos oriundos das definições sobre tesouro.

6.6 Sistema de conceitos

É apresentado aqui, na forma de um grafo, o sistema de conceitos, baseado nas definições dos autores, de modo a permitir uma melhor visão do significado dos termos selecionados. Alguns termos estão representados como entidades (nós), enquanto que outros estão representados como relacionamentos (arestas). Algumas relações foram adicionadas ao sistema (e.g. é-um) para torná-lo mais completo e inteligível. No grafo relacionado

com os conceitos abordados nas definições sobre ontologia (FIG. 19), o conceito de ontologia foi adicionado e colocado de forma destacada, uma vez que é o conceito que está sendo definido. O mesmo ocorre em relação ao conceito de tesouro em seu sistema de conceitos (FIG. 20).

Pode-se observar que o núcleo central do sistema de conceitos sobre ontologia envolve os conceitos de “representação” e de “sistema conceitual”. Propõe-se aqui que uma definição sobre a ontologia na CC deve relacionar os conceitos que formam o núcleo central do sistema de conceitos. Já em relação ao sistema de conceitos para o tesouro pode-se observar que o núcleo central envolve os conceitos relacionados com "vocabulários" e "linguagens".

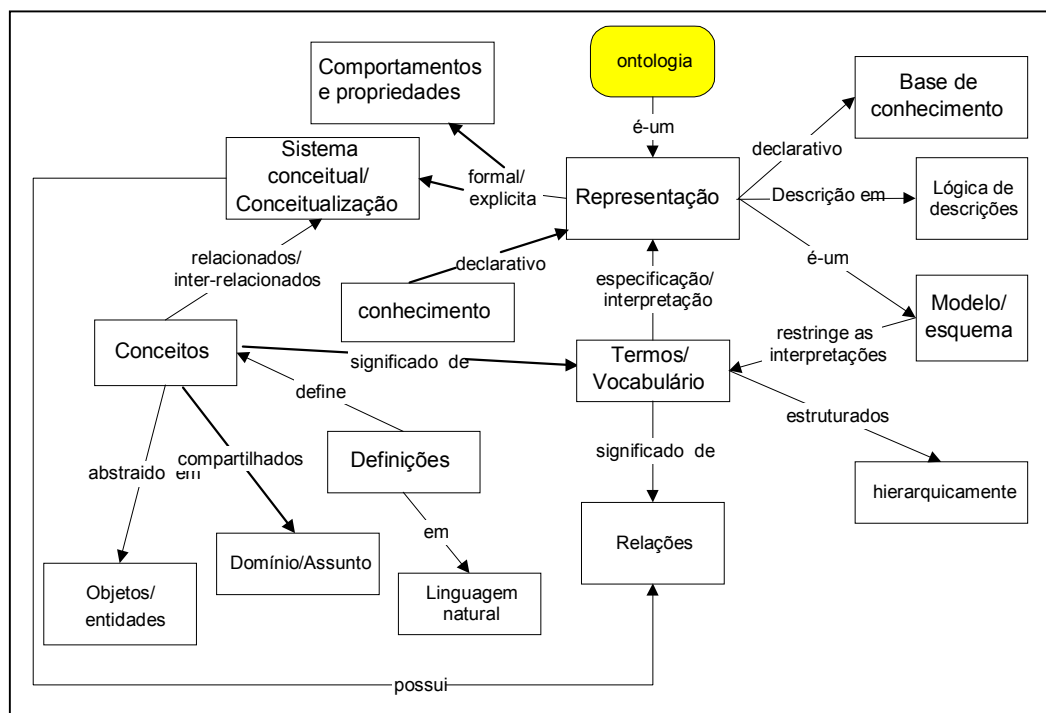


FIGURA 19 – Sistema de conceitos para ontologia

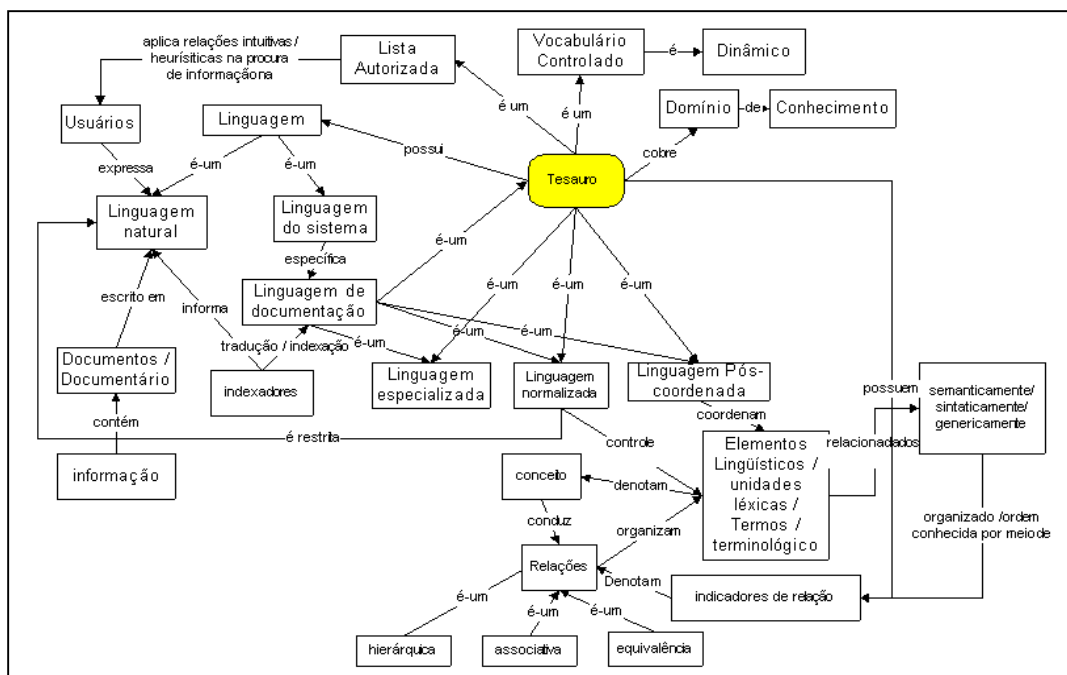


FIGURA 20 – Sistema de conceitos para tesouro

6.7 Análise sobre a ocorrência dos termos nas definições sobre ontologias

Em relação à frequência de ocorrência dos termos pode-se notar que os termos que ocorrem com maior frequência são os termos “domínio”, “termo” e “relacionado/relações” sendo que o primeiro ocorre em sete definições, o segundo ocorre em seis definições e o terceiro ocorre em cinco definições. A ocorrência do primeiro indica que a ontologia na CC está relacionada com domínios específicos e não com o conhecimento independente de domínio, como é o caso da ontologia na filosofia. Já a frequência de ocorrência do termo “termo” indica que em muitos casos a ontologia na CC é uma tentativa de estabelecer uma terminologia para um determinado domínio. A frequência de ocorrência do termo “relacionado/relações” indica que este conceito é um elemento importante na identificação de ontologias. Os demais termos ocorrem com menor frequência, variando de uma a três ocorrências, o que sugere uma falta de uniformidade entre as definições de ontologia no âmbito da ciência da computação.

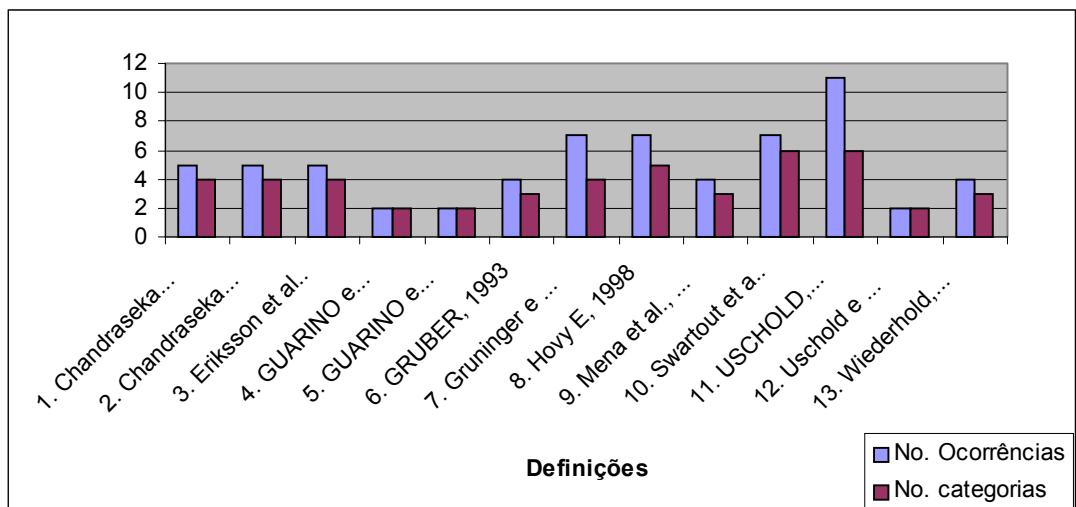


GRÁFICO 1 - Ocorrência de termos e categorias por definição de ontologia

O GRAF. 1 mostra o número de ocorrências de termos e o número de categorias abordadas por cada definição. A definição com a maior ocorrência de termos foi a definição 11 (Uschold, 1996), com a ocorrência de onze termos distintos. Seguida pelas definições 7 (Gruninger e Fox, 1995), 8 (HOVY, 1998) e 10 (Swartout et al. 1996), com a ocorrência de sete termos distintos cada uma. A definição número 6 (Gruber, 1993), que vem sendo adotada como padrão dentro da CC ficou em oitavo lugar, com apenas quatro ocorrências. As definições com maior número de palavras tenderam a ter uma maior ocorrência de termos. De fato, a definição 11 aparenta ser a mais completa, abordando todos os aspectos relacionados com o conceito de ontologia na CC. No entanto, a popularidade da definição 6, pode, em princípio, ser explicada pelo fato de ser, ao mesmo tempo, sucinta e abordar o núcleo central do sistema de conceitos relacionados com a ontologia na CC. Pode-se notar, também, que o número de categorias abordadas pelas definições acompanha a frequência de termos. Portanto, as definições com maior ocorrência de termos também abrangem um número maior de categorias.

O GRAF. 2 mostra o número de ocorrências de termos distribuídos por categoria. Nota-se uma distribuição razoavelmente equilibrada da frequência ocorrência entre os termos, com um certo destaque para a categoria

"atributo" e a ausência de ocorrência na categoria agentes. A ausência de ocorrências pode ser atribuída a uma maior preocupação, da CC, no sentido de armazenagem e processamento por dispositivos computacionais e não no sentido de manipulação por um usuário.

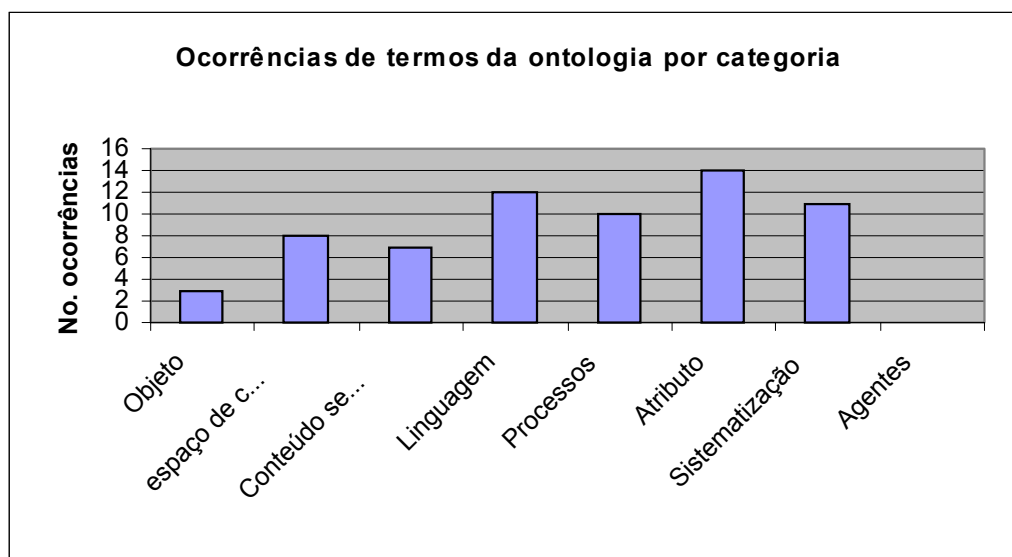


GRÁFICO 2 - Ocorrências de termos de ontologia por categoria

6.8 Análise sobre a ocorrência dos termos nas definições sobre tesauro

Em relação ao número de conceitos, observa-se que, apesar de em menor número, as definições sobre tesauro envolvem quase o mesmo número de termos que as definições sobre ontologia. Já em relação à frequência de ocorrência dos termos pode-se notar que os termos "relacionados/relação" e "termo". O primeiro ocorre em quatro definições e o segundo ocorre em três definições, sendo que em outras definições são encontrados termos equivalentes ou próximos, como "vocabulário" e "léxico". A ocorrência destes termos mostram que os tesauros focalizam no registro de termos e suas relações e qualquer definição sobre este instrumento deve mencionar este foco. Pode-se atribuir a existência de uma definição que não aborda o aspecto relacional dos tesauros (número 6, UNESCO, 1973), ao fato desta

definição descrever os tesouros sob foco do uso e não de suas características intrínsecas.

O GRAF.3 mostra o número de ocorrências de termos e o número de categorias abordadas por cada definição. Em relação à frequência de ocorrência de termos nas definições pode-se observar que existe um certo equilíbrio entre as definições. Uma definição contabilizou doze termos e três definições contabilizaram uma frequência de nove termos, enquanto que duas definições ficaram em um patamar mais abaixo, com seis e quatro ocorrências. As definições que obtiveram uma maior frequência de ocorrência termos foram as de número 1 (ANSI/NISO Z39-19-1993), 2 (Currás, 1995) e 5 (UNESCO, 1973), todas com nove ocorrências e a definição 6 (UNESCO, 1973), com doze ocorrências. Nota-se que três das quatro definições com maior frequência foram criadas por organismos normalizadores. Isto pode explicar a maior abrangência dessas definições, uma vez que organismos normalizadores tendem a emitir definições mais completas. Confirmando esta tendência, pode-se notar que as duas definições com menos ocorrências foram propostas por indivíduos: definição número 3 (Howerton,1965), com seis ocorrências e definição 4 (Alan Gilchrist, 1971) com quatro ocorrências. Outro aspecto importante é o fato das duas definições com as menores frequências de termos serem também as mais antigas. Ou seja, as definições mais recentes tendem a ser mais completas pois se beneficiam do amadurecimento do conceito sobre o instrumento. Isto explica o porque da definição número 2 ser tão completa, apesar de ser proposta por um indivíduo. Pode-se notar, também, que o número de categorias abordadas pelas definições nem sempre acompanha a frequência de termos. A definição 1 abrange um número menor de categorias do que a definição 3, apesar de possuir uma frequência de ocorrência de termos bem maior que a segunda. Adotando-se o critério de que a abrangência de categorias indica uma definição mais abrangente pode-se concluir que a definição 6 é a mais abrangente, seguida pelas definições 2, 3 e 5.

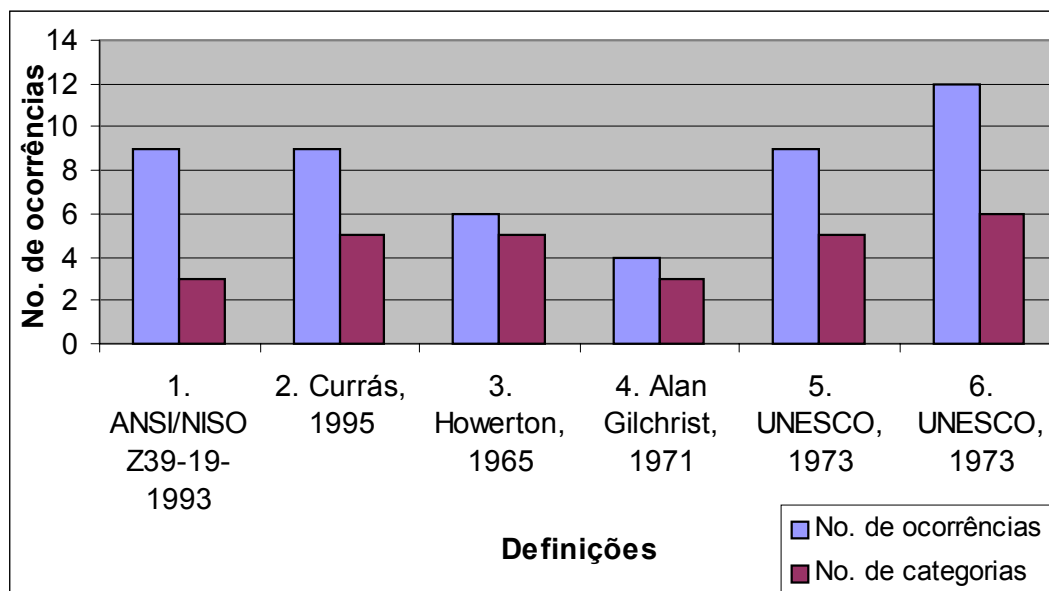


GRÁFICO 3 - Ocorrência de termos e categorias por definição de tesouro

O GRAF. 4 mostra o número de ocorrências de termos distribuídos por categoria. Nota-se que houve um destaque para as categorias atributo e linguagem. O destaque para a categoria linguagem pode ser interpretado com uma maior preocupação, por parte da CI, com a comunicação, uma vez que os tesouros são instrumentos usados na tradução da linguagem natural para uma linguagem de documentação. Já o destaque para a categoria atributo é natural pois é uma categoria que tende a possuir um maior número de termos, uma vez que agrupa os termos que denotam as características (*definiens*) dos objetos que estão sendo definidos, sendo a quantidade destes, função da profundidade da intensão do conceito que está sendo definido. Os termos 'controlada', 'normalizada' e 'autorizada' evidenciam o caráter terminológico dos tesouros.

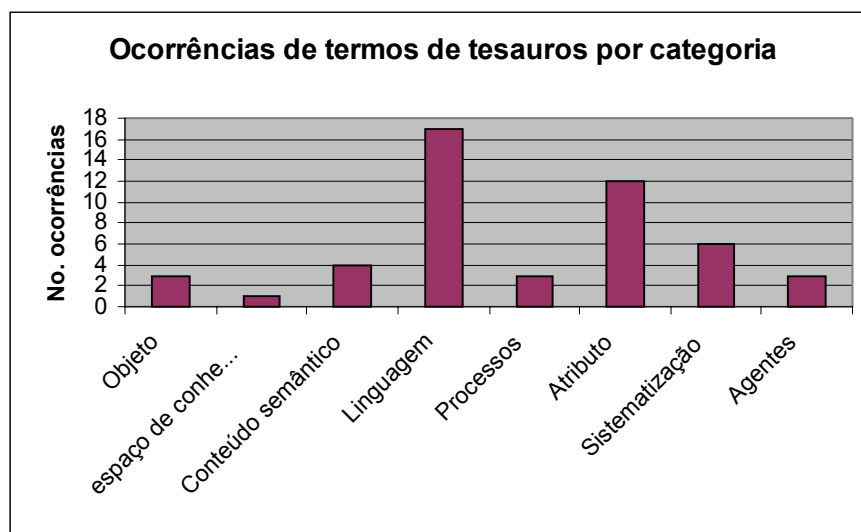


GRÁFICO 4 - Ocorrências de termos de tesouro por categoria

6.9 Análise comparativa sobre a ocorrência dos termos nas definições sobre ontologia e tesouro

Pode-se observar uma certa preocupação com o usuário (termos 'usuário' e 'usuários'), evidenciando a relação do tesouro com sistemas de classificação e recuperação de documentos. A não ocorrência de termos semelhantes nas definições de ontologias pode indicar a diferença de propósitos entre os dois instrumentos. Da mesma forma, a ênfase das definições sobre tesouro na categoria linguagem mostra que os tesouros focam na comunicação entre usuários e sistemas de classificação.

No campo das ontologias os termos 'formais', 'formal' e 'Lógica' (de descrições) revelam uma certa preocupação com o rigor da representação, o que é compreensível, uma vez que a ontologia é usada para trocas informacionais entre sistemas de computação e, portanto, para que não ocorram computações equivocadas, é necessária uma especificação clara das descrições. Já no caso dos tesouros também pressupõem-se que não existe ambigüidades em relação ao significado dos termos; no entanto, esta condição é garantida pela metodologia de construção.

A ontologia privilegia a formalização das propriedades das relações devido à necessidade de realizar inferências por meio de um computador, enquanto que nos tesouros conta-se com interventores humanos para a realização dessas inferências.

7 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O resultado da análise quantitativa deu os subsídios para a análise qualitativa sobre tesouros e ontologias. Obteve-se os seguintes resultados na análise quantitativa:

- categoria "objeto" - nenhuma ocorrência do mesmo termo tanto nas definições de tesouros quanto de ontologias.
- categoria "linguagem" - os termos "linguagem natural" e "termos" são oriundos dos dois contextos. Alguns termos são próximos, como é o caso de "vocabulário" do contexto de ontologias e "vocabulário controlado", no caso de tesouros.
- categoria "espaço de conhecimento" - o termo "domínio" é usado em ambos os contextos.
- categoria "conteúdo semântico" - os termos "conceito" e "conhecimento" são usados em ambos os contextos e os termos "significados" (ontologia) e "semântica" (tesauro) possuem o mesmo sentido.
- categoria "processos" - cada termo origina-se do conjunto de definições de tesouros ou de ontologias mas não dos dois conjuntos. Isto provavelmente se deve à diferença de atividades entre as duas áreas.
- categoria "atributo" abriga a maior quantidade de termos; no entanto, apenas dois termos ocorreram tanto nas definições de tesauro como nas de ontologia.
- categoria "sistematização" - o termo "relação/relações/relacionado(s)" é usado em ambos os contextos.
- categoria "agentes" não registrou ocorrência de termo oriundo do contexto de ontologias.

Desta forma, a partir destas ocorrências buscou-se analisar qualitativamente os conceitos de ontologia e tesauro, utilizando-se dos

subsídios dos dados quantitativos e usando-se fontes bibliográficas do capítulo "Fundamentação teórica". No início deste trabalho foi lançada uma pergunta que sintetizava a motivação e o objetivo deste trabalho: seriam tesouros e ontologias termos que denotam o mesmo instrumento ou denotariam objetos diferentes e, portanto, seu uso indistinto seria inadequado? Acredita-se que a análise empírica realizada sobre as definições, assim como a literatura levantada tenham fornecido subsídios para responder esta questão. A discussão que se segue analisa as impressões coletadas.

Os tesouros da ciência da informação e as ontologias da ciência da computação possuem origens e propósitos distintos. O primeiro nasceu como um instrumento prático para auxiliar na indexação e busca de documentos e o segundo da necessidade de descrever os objetos digitais e suas relações. Pode-se dizer que existem alguns pontos de contato entre essas origens, uma vez que estão relacionadas com a descrição de alguma entidade: assunto de um documento no primeiro caso e objetos e relações no segundo. No entanto, as diferenças também deixaram suas marcas, influenciando na forma final do instrumento de cada vertente. Na ciência da computação a situação é um pouco mais indefinida. Aparentemente tudo que modela um segmento da realidade pode ser denominado de ontologia, uma vez que é a palavra da moda. Neste caso, até mesmo os tesouros podem ser enquadrados como ontologias terminológicas.

A análise quantitativa evidenciou a diferença de propósitos entre os dois instrumentos. A frequência de ocorrência de termos, bem como a abrangência das definições sobre as categorias mostrou que os tesouros têm como propósito, servir de instrumento de registro terminológico e para ser usado por pessoas, e não para registro do conhecimento para inferências computacionais. Já no caso das definições sobre ontologia, a ocorrência de termos tais como "formal" e "Lógica", demonstra a necessidade de registro do conhecimento do domínio em uma linguagem que possa ser processada pelo computador para a realização de inferências.

Outra evidência para esta conclusão é o fato de que, usando linguagens para ontologias (e.g. OIL) é mais fácil registrar certas propriedades das relações do que com o uso de tesouros. No entanto, esta diferença de expressividade não é tão significativa para a tarefa de indexação ou busca de documentos.

A ontologia como um sistema de categorias, tal como é vista na filosofia, ocorre na ciência da informação durante a elaboração de um conjunto de categorias que serão utilizados para organizar sistemas de classificação e recuperação de informação, como pode ser observado nas palavras de Dahlberg:

“Podemos ver que as categorias têm uma capacidade de estrutura: não apenas estruturam, de fato, todos os nossos elementos de conhecimento e unidades do conhecimento; elas fornecem, ao mesmo tempo, por este meio, o esqueleto, os ossos e tendões para estruturar todo o nosso conhecimento. Com seu uso consciencioso, então, o corpo do nosso conhecimento pode se manter unido, pode se mover, pode se manter flexível - e pode crescer organicamente” (Dahlberg, 1978c. p. 34).

Já na ciência da computação este ponto de vista não é adotado. Alguns pesquisadores declaram esta distinção explicitamente (Valente, 1995 e Guarino, 1998). Por exemplo em um de seus artigos Guarino declara:

“No sentido filosófico, podemos nos referir a uma "ontologia" como um sistema particular de categorias que versa sobre uma certa visão do mundo. Desta forma, este sistema não depende de uma linguagem particular: a ontologia de Aristóteles é sempre a mesma, independente da linguagem usada para descrevê-la. Por outro lado, em seu uso mais prevalente na IA, uma ontologia é referida como um artefato de engenharia, constituído de um vocabulário específico usado para descrever uma certa realidade e um conjunto de pressupostos explícitos relacionados com o significado pretendido para as palavras do vocabulário.” (Guarino, 1998b, p.2) (tradução da autora)

Uma ontologia como vista pela ciência da computação é um sistema de conceitos, da mesma forma que os tesouros, e como tal pertence ao nível epistemológico e não ao ontológico. A diferença em relação aos tesouros pode ocorrer em termos de linguagem, de nível de formalização e de propósitos. Neste sentido pode ser adequado que, no âmbito da ciência da

computação, os tesouros sejam enquadrados como ontologias. Contudo, sugere-se que este enquadramento não seja adotado na ciência da informação, uma vez que a ontologia, no sentido filosófico de um sistema de categorias, é utilizada no escopo da CI.

Alguns pesquisadores concordam com a semelhança entre os tesouros e as ontologias da CC, mas alegam que a distinção entre eles estaria no fato das ontologias permitirem uma maior variedade de relações. Tal visão não procede e advém da falta de entendimento do que é um termo e o que é relação segundo a teoria dos tesouros. Os tesouros, assim como algumas linguagens para representação de ontologias, apresentam um conjunto de relações pré-definidas para serem usadas para a estruturação dos conceitos. Este conjunto de relações de estruturação varia de tesouro para tesouro, em função da teoria subjacente e dos propósitos almejados. Já as relações observadas no domínio são representadas nos tesouros da mesma forma que qualquer outro conceito, enquanto que nas ontologias da ciência da computação, as relações são representadas de forma distinta das propriedades (isto é classes) e a elas podem ser atribuídas restrições e propriedades estruturais (e.g. transitividade) que podem ser usadas na realização de inferências.

Devido às pressões para o desenvolvimento de aplicações ou devido à necessidade de um aprofundamento nos estudos em várias áreas de conhecimento (como a filosofia, a lingüística e a ciência da informação), a comunidade de IA tem relutado em desenvolver metodologias que levem em consideração estes vários aspectos.

As conclusões deste trabalho, suportadas pela análise da literatura e pela análise quantitativa, são as seguintes: 1) ontologia da filosofia e ontologia da CC são objetos distintos; 2) ontologia da CC e tesouro são objetos que operam no mesmo nível, ou seja, nível epistemológico; 3) ontologia e tesouro possuem propósitos distintos, sendo que a primeira é voltada para registro de conceitos de um domínio com o propósito de inferência automatizada e o segundo para comunicação entre o usuário e linguagens documentárias; 4) os tesouros cumprem parte dos objetivos que a CC

pretende com o uso de ontologias e por isso são denominados de ontologias terminológicas.

Como trabalhos futuros sugere-se a análise formal da diferença de expressividade entre as linguagens usadas para representação de ontologias e as usadas na elaboração de tesouros e a verificação se a incorporação dos elementos ausentes em um e no outro instrumento permite um melhor cumprimento de seus propósitos. Sugere-se também, em vista da ausência de uma metodologia amplamente adotada para a elaboração de ontologias, a adaptação das metodologias utilizadas na ciência da informação, particularmente a teoria do conceito, para ser usada como base para a construção de ontologias. Finalmente, sugere-se a aplicação do método analítico-sintético como forma de análise de conteúdo para outros temas da área.

8 ÍNDICE

A

análise de conteúdo, 11, 13, 23, 60, 61, 100

C

ciência da informação, 3, 10, 13, 15, 16, 18, 23, 24, 34, 60, 62, 63, 64, 68, 98, 99, 100
conhecimento, 3, 10, 12, 13, 17, 19, 21, 22, 24, 30, 39, 40, 41, 42, 43, 44, 45, 46, 48, 49, 50, 51, 52, 53, 54, 55, 58, 60, 61, 64, 67, 71, 72, 74, 75, 76, 78, 80, 86, 89, 96, 97, 98, 99

E

epistemológico, 3, 50, 53, 98, 99

F

filosofia, 15, 30, 31, 34, 37, 38, 39, 58, 60, 67, 89, 98, 99

I

informação, 10, 12, 13, 15, 17, 24, 34, 40, 41, 45, 51, 71, 75, 78, 98
inteligência artificial, 13, 15, 17, 30, 38, 39, 58, 60, 71

K

KeyWord Out of Context
KeyWord Out of Context, 69

L

linguagem natural, 10, 24, 41, 44, 47, 52, 73, 80, 85, 93, 96
Lógica de Primeira Ordem, 53

M

método analítico-sintético, 2, 3, 14, 15, 16, 60, 62, 100

O

ontologia, 3, 8, 10, 13, 15, 30, 31, 32, 33, 34, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 57, 58, 59, 62, 63, 64, 67, 68, 74, 86, 87, 88, 89, 90, 91, 94, 95, 96, 97, 98, 99
ontologias, 10, 11, 13, 15, 16, 23, 28, 30, 41, 42, 43, 47, 48, 54, 59, 60, 61, 63, 64, 65, 67, 68, 72, 84, 85, 89, 94, 96, 97, 98, 99, 100
ontologia informal, 14

T

terminológico, 10, 13, 24, 60, 93, 97
termos, 3, 12, 13, 16, 17, 18, 21, 23, 24, 25, 27, 28, 29, 30, 36, 37, 41, 44, 46, 47, 48, 53, 62, 63, 69, 70, 71, 72, 73, 75, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, 90, 91, 92, 93, 94, 96, 97, 98
tesauro, 3, 13, 14, 15, 23, 24, 25, 27, 28, 29, 30, 37, 47, 48, 57, 62, 63, 64, 68, 81, 86, 87, 88, 89, 91, 93, 94, 96, 99, 1, 9
tesauros, 10, 11, 13, 15, 16, 23, 24, 25, 26, 27, 28, 29, 30, 37, 39, 48, 56, 57, 60, 61, 62, 63, 64, 68, 71, 72, 78, 84, 85, 86, 93, 94, 95, 96, 97, 98, 99, 100

9 BIBLIOGRAFIA

1. ALVARENGA, Lídia. Definição de publicações oficiais. R. *Esc. Bibliotecon. UFMG*, Belo Horizonte, v. 22, n. 2, p. 213-238, Jul/Dez.1993.
2. ANDER-EGG, Ezequiel. *Introducción a las técnicas de investigación social: para trabajadores sociales*. 7.ed. Buenos Aires : Hmanitas, 1978.
3. ANSI Z39.19:1993. *American National Standard Guidelines for Thesaurus Structure, Construction, and Use*. New York, 1993
4. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS , Rio de Janeiro. *Normas ABNT sobre documentação*. Rio de Janeiro, 1989. (Coletânea de normas).
5. BRACHMAN, R. J. On the epistemological status of semantic networks. In: N.V. Findler (Ed.) *Associative Networks: Representation and use of knowledge by computers*. Academic Press : [s.l.], 1979. p. 3-50.
6. BRAGA, José L., LAENDER, Alberto H. F., RAMOS, Claudiney V. A knowledge-based approach to cooperative relational database querying. *International Journal of Pattern Recognition and Artificial Intelligence*, v.1, n.14, p. 73-90, Feb. 2000.
7. BUNGE, M. *Ontology I: the furniture of the world*. Dordrecht, holland : D. Reidel, 1977. (Treatise on basic Philosophy, v. 3).
8. CAMPOS, M. L. A. *Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia*. Rio de Janeiro : Escola de comunicação UFRJ/IBICT, 1974. 190 p. (Dissertação, Mestrado em Ciência da Informação).
9. _____. *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói : Eduff, 2001a.
10. _____. *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria*. Rio de Janeiro : Escola de comunicação UFRJ/IBICT, 2001b.(Tese em Ciência da Informação).
11. CHANDRASEKARAN, B., JOSEPHSON, J. R., BENJAMINS, V. R. What are Ontologies, and why do we need them. *IEEE Intelligent Systems*, v. 14, n. 1, p.20-26, Jan./Feb. 1999.

12. CINTRA, Anna Maria Marques *et al.* *Para entender as linguagens documentárias*. 2. ed. rev. ampl. São Paulo : Polis, 2002.
13. CRANFIELD S., PURVIS, M., NOWOSTAWSKI, M. *Is it an ontology or an abstract syntax? modelling objects, knowledge and agent messages*. In: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 14, 2000, [s.l.]. *Proceedings...* [s.l.], 2000. p. 16.1-16.4.
14. CURRÁS, Emilia. *Tesauros: linguagens terminológicas*. Brasília : IBICT, 1995. 286 p.
15. DAHLBERG, Ingetraut. Teoria do conceito. *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978a.
16. _____. A Referent-Oriented, Analytical Concept Theory for INTERCONCEPT. *International Classification*, v. 5, n. 3, p. 142-151, 1978b.
17. _____. *Ontical structures and universal classification*. Bangalore : Sarada Ranganthan Endowment, 1978c. 64 p.
18. _____. Conceptual definitions for INTERCONCEPT. *International Classification*, v. 8, n. 1, p. 16-22, 1981.
19. DODEBEI, Vera Lúcia Doyle. *Tesauro: linguagem de representação da memória documentária*. Niterói: Intertexto; Rio de Janeiro : Interciência, 2002.
20. DUARTE, M. M. *Ferramenta de Apoio à Identificação de Fatores Críticos de Sucesso em uma Organização*. Belo Horizonte : Universidade Federal de Minas Gerais, Ago. 2002. (Dissertação, Mestrado em Ciência da Computação).
21. DUINEVELD, A. J., STOTER, R., WEIDEN, M. R., KENEP, B., BENJAMINS, V. R. WonderTools? A comparative study of ontological engineering tools. In: WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT VOYAGER, 20, Saturday 16 to Thursday 21 October, 1999, Alberta, Canada. *Anais...* Alberta, 1999.
22. ENDERTON, H. B. A. *Mathematical Introduction to Logic*. [s.l.] : Academic Press, 1972.
23. ERDMANN, Michael, STUDER, Rudi. Ontologies as Conceptual Models for XML Documents. In: WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT VOYAGER, 20, Saturday 16 to Thursday 21 October, 1999, Alberta, Canada. *Anais...* Alberta, 1999.

24. ERIKSSON, Henrik, PUERTA, Angel R., MUSEN, Mark A. Generation of knowledge acquisition tools from domain ontologies. In: SYSTEMS WORKSHOP, BANFF KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED, 8, 1994, Banff, Canada. v. 1, p. 7-20, Banff, Canada, 1994.
25. FRANÇA, Júlia Lessa. *Manual para normalização de publicações técnico-científicas*. 4.ed. rev. e aum. Belo Horizonte : UFMG, 1999. 213 p.
26. FELBER, H., NEDOBITY, W. The basis of Wüster's classification research. In: INT. STUDY CONF. ON CLASSIFICATION RESEARCH OF FID/CR, 4, 28 June-2 Jul. 1982, Augsburg. *Anais...* Augsburg, 1982. p. 95-101.
27. FELBER, H. *Terminology Manual*. Paris : UNESCO, 1984. 234 p. p. 17.
28. FENSEL, D. *et al.* "OIL in a nutshell. In: EUROPEAN KNOWLEDGE ACQUISITION CONFERENCE (EKAW 2000), 2000, Berlin. DIENG, R. *et al.* (Eds). *Lecture Notes in Artificial Intelligence...* n.1937. Berlin : Springer-Verlag, 2000. p.1-16.
29. _____. OIL: an ontology infrastructure for the semantic web. *IEEE Intelligent Systems*, p. 38-45, Mar./Apr. 2001.
30. FERNÁNDEZ, M., GÓMEZ-PÉREZ, A., JURISTO, N. Methontology: from ontological art towards ontological engineering. In: WORKSHOP ON ONTOLOGICAL ENGINEERING (AAAI - Spring Symposium Series), 1997, USA. *Proceedings...USA* : Stanford, 1997.
31. FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro : Nova Fronteira, 1975.
32. FOSKETT, D. *Subject and information analysis*. New York : Marcel Dekker, 1985. p. 270-316: Thesaurus.
33. GARCÍA MORENTE, M. *Fundamentos da filosofia: lições preliminares*. São Paulo : Mestre Jou, 1964.
34. GENESERETH, M. R., FIKES, R. E. *Knowledge Interchange Format: version 3.0, reference manual*. Stanford : Computer Science Dept. Stanford University, 1992. (Relatório Logic-92-1).
35. GILCHRIST, Alan. *The thesaurus in retrieval*. London : ASLIB, 1971.
36. GLINSKI, Wieslaw. Na intelligent quasi-thesaurus for accessing

- information systems. *Transinformação* , v. 9, n. 2, p. 1-9, maio/ago. 1997.
37. GOMES, H. E., CAMPOS, M. L. A. Classificação e Terminologia. In: SIMPÓSIO LATINO AMERICANO DE TERMINOLOGIA, 2, 1990, Brasília. *Anais...* Brasília : IBICT, 1992. p. 252-258.
 38. _____, et al. *Manual de elaboração de tesauros monolíngües*. Brasília : CNPq / PNBU, 1990. 78 p.
 39. GRUBER, T. R. *Ontolingua: a mechanism to support portable ontologies*. Stanford : *Knowledge Systems Laboratory* Stanford University, 1992. (Relatório KSL91-66).
 40. GRUNINGER, M, FOX, M. S. Methodology for the design and evaluation of ontologies. In INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1995, *Montreal, Quebec, Canada*. *Proceedings...* Montreal, Quebec, Canada, 1995.
 41. GUARINO, Nicola. The Ontological Level. In: CASATI, R., SMITH, B., WHITE, G. (Eds.). *Philosophy and the Cognitive Science*. Vienna : Holder-Pichler-Tempsky, 1994. p. 443-456.
 42. GUARINO, Nicola, GIARETTA, P. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: N. MARS (Ed.) *Towards very large Knowledge Bases: Knowledge Building and Knowledge sharing*, 1995a. Amsterdam : IOS Press. p. 25-32.
 43. GUARINO, Nicola. Formal Ontology, conceptual analysis and knowledge representation. GUARINO, Nicola, POLI, Roberto (Eds). *International Journal of Human-Computer Studies*, v. 43, n. 5/6, 1995b. Número especial.
 44. _____. Formal ontology and information systems. In: FOIS, 1998, Trento. *Proceedings...* Amsterdam : IOS Press, 1998. p. 3-15.
 45. GUARINO, Nicola, MASDOLO, Claudio, VETERE, Guido. Ontoseek: content-based access to the Web. *IEEE Intelligent Systems*, p. 70-80, May/June 1999.
 46. GUARINO, Nicola, WELTY, Christopher. Towards a methodology for ontology based model engineering. In: ECOOP-2000 WORKSHOP ON MODEL ENGINEERING, 2000, Cannes, France. *Proceedings...* Cannes, France , 2000a.
 47. _____, A Formal Ontology of Properties. In DIENG, R., CORBY, O. (Eds.). In: INTERNATIONAL CONFERENCE KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT: METHODS,

- MODELS AND TOOLS (EKAW2000), 12, 2000. *Conference...[s.l.]* : Springer Verlag, 2000b. p. 97-112.
48. _____, Identity and subsumption. In GREEN, R., BEAN, C. A., HYON MYAENG, S. (Eds.). *The semantics of relationships: an interdisciplinary perspective*. Kluwer, 2001.
49. _____, Evaluating Ontological Decisions With Ontoclean. *Communications of the Acm*, v. 45, n. 2, Feb. 2002.
50. HOVY, E. Combining and standardizing large-scale, practical ontologies for machine translation and other uses. In: LANGUAGE RESOURCES AND EVALUATION CONFERENCE, May 1998, Granada, Spain. *Conferências...* Granada, Spain, May 1998.
51. HOWERTON, Paul. Organic and functional concepts of authority files. In: NEWMAN, Simon M. (Ed). *Information Systems compatibility*. Washington : Spartan Books, 1965.
52. ISO 1087-1990. *Terminology-Vocabulary*. 27 p.
53. JASPER R., USCHOLD, M. A. Framework for understanding and classifying ontology applications. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE (IJCAI-99), July 31 - August 6, 1999, Stockholm, Sweden. *Proceedings...* Stockholm, Sweden, July 31 - August 6, 1999.
54. KEIZER, N. F. de , ABU-HANNA, A., ZWETSLOOT-SCHONK, J. H. M. Understanding terminological systems I: terminology and typology. *Methods of information in medicine*, v. 39, p. 16–21, 2000.
55. LANCASTER, F. W. *Construção e uso de tesouros: curso condensado*. Brasília : IBICT, 1987. 114 p.
56. LANCASTER, F. W. *Vocabulary control for information retrieval*. 2. ed. Virgínia : IRP, 1986. 270 p.
57. LAWRENCE, S. Access to Scientific Literature. London, In: BUTLER, Declan (Ed.). *The Nature Yearbook of Science and Technology*. England : Macmillan, 2001. p. 86–88.
58. LAWRENCE, Steve S., BOLLACKER K., LEE GILES, C. *Digital Libraries and Autonomous Citation Indexing*. *IEEE Computer*, v. 32, n. 6, p. 67-71, 1999.
59. LENAT, D. B., GUHA, R. V. *Building large knowledge-based systems: representation and inference in the CYC project*. Massachusetts :

Addison-Wesley, 1990.

60. LETHBRIDGE, T. C. *Practical Techniques for Organizing and Measuring Knowledge*. Ottawa : 1994. (Thesis Presented to the School of Graduate Studies and Research. University of Ottawa).
61. MARCONI, Marina de Andrade, LAKATOS, Eva Maria. *Técnicas de pesquisa: planejamento e execução de pesquisas, amostragens e técnicas de pesquisa, elaboração, análise e interpretação de dados*. São Paulo : Atlas, 1982.
62. MENA, E., KASHYAP, V., SHETH, A. *et al.* OBSERVER: an approach for query processing in Global Information systems based on Interoperation across pre-existing Ontologies. In: IFCIS - INTERNATIONAL CONFERENCE ON COOPERATIVE INFORMATION SYSTEMS (COOPIS '96), 1996. *Proceedings...* Brussels, Belgium, June 1996.
63. MICHAELIS. Dicionário: versão eletrônica.
64. MILLER, George. Wordnet: a lexical database for english. *Comm. ACM*, vol. 38, n. 11, p. 39-41, nov. 1995.
65. MORA, J. F. On the early history of Ontology. *Philosophy and Phenomenological Research*, [s.l.] , v. 24, p. 36-47, 1963.
66. MOTTA, Dilza Fonseca da. *Método relacional como nova abordagem para a construção de tesouros*. Rio de Janeiro: Senai, 1987. p. 89. (Coleção Albano Franco, 12).
67. NARDI, D.; BRACHMAN, R. J. An Introduction to Description Logics. In: BAADER, F., CALVANESE, D., MCGUINNESS, D. L., NARDI, D., PATEL-SCHNEIDER, P.F. (Ed.). *In the Description Logic Handbook*. Cambridge : University Press, 2002. p. 5-44.
68. NEBOBITY, W. Terminology and artificial intelligence. *International Classification*, v. 12, n. 1, p. 17-19, 1995.
69. NOCETTI, Milton A., FIGUEIREDO, Regina Célia. Línguas naturais e linguagens documentárias: traços inerentes e ocorrências de interação. *R. Bibliotecon.*, Brasília, v. 6, n. 1, p. 23-37, jan./jun.1978.
70. 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*. Trad. de Álvaro Cabral. Rio : Zahar, 1972. 349 p.
71. OLIVEIRA, A. P. *Ferramentas para estruturação do conhecimento*. Rio de Janeiro : IME, 1990. 106 p. (Dissertação, Mestrado em Sistemas de

Computação).

72. OS PENSADORES. Aristóteles. São Paulo : Nova Cultural, 2000.
73. PAIM, Isis. *Notas de aula*. Belo Horizonte : [s.n.], 199?.
74. POLI, R. *Framing ontology*. 2001. Available from World Wide Web: [http://www.formalontology.it/Framing first.htm](http://www.formalontology.it/Framing%20first.htm).
75. QUINE, W. V. *The Philosophy of Logic*. 2nd ed. [s.l.] : Harvard Univ. Pr, June 1986.
76. PONTES, F. V. *Gestão de Conhecimento Apoiada por Ontologias*. Belo Horizonte : Universidade Federal de Minas Gerais, mar. 2002. (Dissertação, Mestrado em Ciência da Computação).
77. RAMOS, Claudiney V., BRAGA, José L., LAENDER, Alberto H. F. Cooperative querying in relational databases. In: PROCEEDINGS OF INTERNATIONAL CONFERENCE OF THE CHILEAN COMPUTER SCIENCE SOCIETY, 17, Nov. 1997, Valparaiso, Chile. *Prodeedings...* Valparaiso, Chile : IEEE COMPUTER SOCIETY, Nov. 1997. p.190-198.
78. RICH, E., KNIGHT, K. *Inteligência Artificial*. 2.ed. [s.l.] : McGraw-Hill, 1993.
79. RODEAU, G, FELBER, F. L'étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et les sciences des choses. In: TEXTES CHOISIS DE TERMINOLOGIE. I. FONDEMENTS THÉORIQUES DE LA TERMINOLOGIE, Québec, 1981. p. 57-114.
80. RUSSELL, B. *História do Pensamento Ocidental: a aventura dos Pré-socráticos a Wittgenstein*. Rio de Janeiro : Ediouro, 2001.
81. RUSSELL, Stuart, NORVIG, Peter. *Artificial intelligence: a modern approach*. New Jersey : Prentice Hall, 1995. 932 p.
82. SAUSSURE, F. de. *Curso de lingüística geral*. São Paulo : Cultrix, 1987. 279 p.
83. SELLTIZ, C. et al. *Métodos de pesquisa nas relações sociais*. São Paulo : Heder, 1965.
84. SOWA, J. F. Top-layer ontological categories. *International journal of human-computer studies*, 43, p. 669-685, 1995.

85. SOWA, J. F. *Knowledge Representation: Logical, Philosophical and Computational Foundations*. California : Brooks/Cole Publishing Co., 2000.
86. SOWA, J. F. *Building, sharing and merging ontologies*. 2001. Available from World Wide Web: <<http://www.jfsowa.com/ontology/ontoshar.htm>>.
87. SWARTOUT, B., PATIL, R., KNIGHT, K., RUSS, T. Toward Distributed Use of Large-Scale Ontologies. In: WORKSHOP KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS, 20, 1996. Banff, Canada, 1996. *Proceedings...* Banff, Canada, 1996.
88. TAMMA, Valentina, BENCH-CAPON, Trevor. A conceptual model to facilitate knowledge sharing in multi-agent systems. In: WORKSHOP ON ONTOLOGIES IN AGENT SYSTEMS, 29 May 2001, Montreal. *Proceedings of the autonomous agents 2001*, Montreal, 2001. p. 69-76.
89. UNESCO. Guidelines for the establishment and development of monolingual thesauri. Paris, 1973. 37 p.
90. USCHOLD, M. Building Ontologies: Towards a Unified Methodology. proceedings of Expert Systems. In: ANNUAL CONFERENCE OF THE BRITISH COMPUTER SOCIETY SPECIALIST GROUP ON EXPERT SYSTEMS, 16, p. 16-18 Dec. 1996, Cambridge. *Anais...* Cambridge, 1996.
91. USCHOLD, M., GRUNINGER, M. Ontologies: principles, methods and applications. *Knowledge Engineering Review*, v. 11, n. 2, June 1996.
92. VALENTE, A. *Legal knowledge engineering: a modelling approach*. Amsterdam : IOS Press, 1995.
93. VICKERY, B. C. *Classificação e indexação nas ciências*. Rio de Janeiro : BNG/ Brasilart, 1980. 274 p.
94. WELTY, C., GUARINO, N. Supporting Ontological Analysis of Taxonomic Relationships. *Data and Knowledge Engineering*, v.39, n.1, p. 51-74, 2001.
95. WIEDERHOLD, Gio. Interoperation, Mediation, and Ontologies. In: INTERNATIONAL SYMPOSIUM ON FIFTH GENERATION COMPUTER SYSTEMS (FGCS), Dec. 1994, Tokyo, Japan. *Proceedings Workshop on heterogeneous, cooperative knowledge-bases*. Tokyo, Japan : ICOT, Dec. 1994. v. W3, p. 33-48.

10 REFERÊNCIA BIBLIOGRÁFICA

1. ALVARENGA, Lídia. Definição de publicações oficiais. R. *Esc. Bibliotecon. UFMG*, Belo Horizonte, v. 22, n. 2, p. 213-238, Jul/Dez.1993.
2. ANDER-EGG, Ezequiel. *Introducción a las técnicas de investigación social: para trabajadores sociales*. 7.ed. Buenos Aires : Hmanitas, 1978.
3. ANSI Z39.19:1993. *American National Standard Guidelines for Thesaurus Structure, Construction, and Use*. New York, 1993
4. BRACHMAN, R. J. On the epistemological status of semantic networks. In: N.V. Findler (Ed.) *Associative Networks: Representation and use of knowledge by computers*. Academic Press : [s.l.], 1979. p. 3-50
5. BRAGA, José L., LAENDER, Alberto H. F., RAMOS, Claudiney V. A knowledge-based approach to cooperative relational database querying. *International Journal of Pattern Recognition and Artificial Intelligence*, v.1, n.14, p. 73-90, Feb. 2000.
6. BUNGE, M. *Ontology I: the furniture of the world*. Dordrecht, holland : D. Reidel, 1977. (Treatise on basic Philosophy, v. 3).
7. CAMPOS, M. L. A. *Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia*. Rio de Janeiro : Escola de comunicação UFRJ/IBICT, 1974. 190 p. (Dissertação, Mestrado em Ciência da Informação).
8. _____. *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói : Eduff, 2001a.
9. _____. *A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria*. Rio de Janeiro : Escola de comunicação UFRJ/IBICT, 2001b.(Tese em Ciência da Informação).
10. CHANDRASEKARAN, B., JOSEPHSON, J. R., BENJAMINS, V. R. What are Ontologies, and why do we need them. *IEEE Intelligent Systems*, v. 14, n. 1, p.20-26, Jan./Feb. 1999.
11. CRANFIELD S., PURVIS, M., NOWOSTAWSKI, M. *Is it an ontology or an abstract syntax? modelling objects, knowledge and agent messages*.

- In: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 14, 2000, [s.l.]. *Proceedings...* [s.l.], 2000. p. 16.1-16.4.
12. CURRÁS, Emilia. *Tesauros: linguagens terminológicas*. Brasília : IBICT, 1995. 286 p.
 13. DAHLBERG, Ingetraut. Teoria do conceito. *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978a.
 14. _____. A Referent-Oriented, Analytical Concept Theory for INTERCONCEPT. *International Classification*, v. 5, n. 3, p. 142-151, 1978b.
 15. _____. *Ontical structures and universal classification*. Bangalore : Sarada Ranganthan Endowment, 1978c. 64 p.
 16. _____. Conceptual definitions for INTERCONCEPT. *International Classification*, v. 8, n. 1, p. 16-22, 1981.
 17. DUARTE, M. M. *Ferramenta de Apoio à Identificação de Fatores Críticos de Sucesso em uma Organização*. Belo Horizonte : Universidade Federal de Minas Gerais, Ago. 2002. (Dissertação, Mestrado em Ciência da Computação).
 18. DUINEVELD, A. J., STOTER, R., WEIDEN, M. R., KENEP, B., BENJAMINS, V. R. WonderTools? A comparative study of ontological engineering tools. In: WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT VOYAGER, 20, Saturday 16 to Thursday 21 October, 1999, Alberta, Canada. *Anais...* Alberta, 1999.
 19. ENDERTON, H. B. A. *Mathematical Introduction to Logic*. [s.l.] : Academic Press, 1972.
 20. ERDMANN, Michael, STUDER, Rudi. Ontologies as Conceptual Models for XML Documents. In: WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT VOYAGER, 20, Saturday 16 to Thursday 21 October, 1999, Alberta, Canada. *Anais...* Alberta, 1999.
 21. ERIKSSON, Henrik, PUERTA, Angel R., MUSEN, Mark A. Generation of knowledgeacquisition tools from domain ontologies. In: SYSTEMS WORKSHOP, BANFF KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED, 8, 1994, Banff, Canada. v. 1, p. 7-20, Banff, Canada, 1994.
 22. FENSEL, D. *et al.* "OIL in a nutshell. In: EUROPEAN KNOWLEDGE ACQUISITION CONFERENCE (EKAW 2000), 2000, Berlin. DIENG, R. *et*

- al. (Eds). *Lecture Notes in Artificial Intelligence...* n.1937. Berlin : Springer-Verlag, 2000. p.1–16.
- 23._____. OIL: an ontology infrastructure for the semantica web. *IEEE Intelligent Systems*, p. 38-45, Mar./Apr. 2001.
- 24.FERNÁNDEZ, M., GÓMEZ-PÉREZ, A., JURISTO, N. Methontology: from ontological art towards ontological engineering. In: WORKSHOP ON ONTOLOGICAL ENGINEERING (AAAI - Spring Symposium Series), 1997, USA. *Proceedings...*USA : Stanford, 1997.
- 25.FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro : Nova Fronteira, 1975.
- 26.FOSKETT, D. *Subject and information analysis*. New York : Marcel Dekker, 1985. p. 270-316: Thesaurus.
- 27.GARCÍA MORENTE, M. *Fundamentos da filosofia: lições preliminares*. São Paulo : Mestre Jou, 1964.
- 28.GILCHRIST, Alan. *The thesaurus in retrieval*. London : ASLIB, 1971.
- 29.GOMES, H. E., et al. *Manual de elaboração de tesauros monolíngües*. Brasília : CNPq / PNB, 1990. 78 p.
- 30._____, CAMPOS, M. L. A. Classificação e Terminologia. In: SIMPÓSIO LATINO AMERICANO DE TERMINOLOGIA, 2, 1990, Brasília. *Anais...* Brasília : IBICT, 1992. p. 252-258.
- 31.GRUBER, T. R. *Ontolingua: a mechanism to support portable ontologies*. Stanford : Knowledge Systems Laboratory Stanford University, 1992. (Relatório KSL91-66).
- 32.GRUNINGER, M, FOX, M. S. Methodology for the design and evaluation of ontologies. In INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1995, Montreal, Quebec, Canada. *Proceedings...* Montreal, Quebec, Canada, 1995.
- 33.GUARINO, Nicola. The Ontological Level. In: CASATI, R., SMITH, B., WHITE, G. (Eds.). *Philosophy and the Cognitive Science*. Vienna : Holder-Pichler-Tempsky, 1994. p. 443-456.
- 34.GUARINO, Nicola, GIARETTA, P. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: N. MARS (Ed.) *Towards very large Knowledge Bases: Knowledge Building and Knowledge sharing*, 1995a. Amsterdam : IOS Press. p. 25-32.

35. GUARINO, Nicola. Formal Ontology, conceptual analysis and knowledge representation. GUARINO, Nicola, POLI, Roberto (Eds). *International Journal of Human-Computer Studies*, v. 43, n. 5/6, 1995b. Número especial.
36. _____. Formal ontology and information systems. In: FOIS, 1998, Trento. *Proceedings...* Amsterdam : IOS Press, 1998. p. 3-15.
37. GUARINO, Nicola, MASDOLO, Claudio, VETERE, Guido. Ontoseek: content-based access to the Web. *IEEE Intelligent Systems*, p. 70-80, May/June 1999.
38. GUARINO, Nicola, WELTY, Christopher. Towards a methodology for ontology based model engineering. In: ECOOP-2000 WORKSHOP ON MODEL ENGINEERING, 2000, Cannes, France. *Proceedings...* Cannes, France , 2000a.
39. _____, A Formal Ontology of Properties. In DIENG, R., CORBY, O. (Eds.). In: INTERNATIONAL CONFERENCE KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT: METHODS, MODELS AND TOOLS (EKAW2000), 12, 2000. *Conference...[s.l.]* : Springer Verlag, 2000b. p. 97-112.
40. _____, Identity and subsumption. In GREEN, R., BEAN, C. A., HYON MYAENG, S. (Eds.). *The semantics of relationships: an interdisciplinary perspective*. Kluwer, 2001.
41. _____, Evaluating Ontological Decisions With Ontoclean. *Communications of the Acm*, v. 45, n. 2, Feb. 2002.
42. HOVY, E. Combining and standardizing large-scale, practical ontologies for machine translation and other uses. In: LANGUAGE RESOURCES AND EVALUATION CONFERENCE, May 1998, Granada, Spain. *Conferências...* Granada, Spain, May 1998.
43. HOWERTON, Paul. Organic and functional concepts of authority files. In: NEWMAN, Simon M. (Ed). *Information Systems compatibility*. Washington : Spartan Books, 1965.
44. ISO 1087-1990. *Terminology-Vocabulary*. 27 p.
45. JASPER R., USCHOLD, M. A. Framework for understanding and classifying ontology applications. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE (IJCAI-99), July 31 - August 6, 1999, Stockholm, Sweden. *Proceedings...* Stockholm, Sweden, July 31 - August 6, 1999.

46. KEIZER, N. F. de , ABU-HANNA, A., ZWETSLOOT-SCHONK, J. H. M. Understanding terminological systems I: terminology and typology. *Methods of information in medicine*, v. 39, p. 16–21, 2000.
47. LANCASTER, F. W. *Construção e uso de tesouros: curso condensado*. Brasília : IBICT, 1987. 114 p.
48. LANCASTER, F. W. *Vocabulary control for information retrieval*. 2. ed. Virgínia : IRP, 1986. 270 p.
49. LAWRENCE, S. Access to Scientific Literature. London, In: BUTLER, Declan (Ed.). *The Nature Yearbook of Science and Technology*. England : Macmillan, 2001. p. 86–88.
50. LAWRENCE, Steve S., BOLLACKER K., LEE GILES, C. *Digital Libraries and Autonomous Citation Indexing*. *IEEE Computer*, v. 32, n. 6, p. 67-71, 1999.
51. LENAT, D. B., GUHA, R. V. *Building large knowledge-based systems: representation and inference in the CYC project*. Massachusetts : Addison-Wesley, 1990.
52. LETHBRIDGE, T. C. *Practical Techniques for Organizing and Measuring Knowledge*. Ottawa : 1994. (Thesis Presented to the School of Graduate Studies and Research. University of Ottawa)
53. MARCONI, Marina de Andrade, LAKATOS, Eva Maria. *Técnicas de pesquisa: planejamento e execução de pesquisas, amostragens e técnicas de pesquisa, elaboração, análise e interpretação de dados*. São Paulo : Atlas, 1982.
54. MENA, E., KASHYAP, V., SHETH, A. *et al.* OBSERVER: an approach for query processing in Global Information systems based on Interoperation across pre-existing Ontologies. In: IFCIS - INTERNATIONAL CONFERENCE ON COOPERATIVE INFORMATION SYSTEMS (COOPIS '96), 1996. *Proceedings...* Brussels, Belgium, June 1996.
55. MICHAELIS. Dicionário: versão eletrônica.
56. MILLER, George. Wordnet: a lexical database for english. *Comm. ACM*, vol. 38, n. 11, p. 39-41, nov. 1995.
57. MORA, J. F. On the early history of Ontology. *Philosophy and Phenomenological Research*, [s.l.] , v. 24, p. 36-47, 1963.

58. NARDI, D.; BRACHMAN, R. J. An Introduction to Description Logics. In: BAADER, F., CALVANESE, D., MCGUINNESS, D. L., NARDI, D., PATEL-SCHNEIDER, P.F. (Ed.). *In the Description Logic Handbook*. Cambridge : University Press, 2002. p. 5-44.
59. NOCETTI, Milton A., FIGUEIREDO, Regina Célia. Línguas naturais e linguagens documentárias: traços inerentes e ocorrências de interação. *R. Bibliotecon.*, Brasília, v. 6, n. 1, p. 23-37, jan./jun.1978.
60. 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*. Trad. de Álvaro Cabral. Rio : Zahar, 1972. 349 p.
61. PAIM, Isis. *Notas de aula*. Belo Horizonte : [s.n.], 199?.
62. POLI, R. *Framing ontology*. 2001. Available from World Wide Web: [http://www.formalontology.it/Framing first.htm](http://www.formalontology.it/Framing%20first.htm).
63. PONTES, F. V. *Gestão de Conhecimento Apoiada por Ontologias*. Belo Horizonte : Universidade Federal de Minas Gerais, mar. 2002. (Dissertação, Mestrado em Ciência da Computação).
64. QUINE, W. V. *The Philosophy of Logic*. 2nd ed. [s.l.] : Harvard Univ. Pr, June 1986.
65. RAMOS, Claudiney V., BRAGA, José L., LAENDER, Alberto H. F. Cooperative querying in relational databases. In: PROCEEDINGS OF INTERNATIONAL CONFERENCE OF THE CHILEAN COMPUTER SCIENCE SOCIETY, 17, Nov. 1997, Valparaiso, Chile. *Prodeedings...* Valparaiso, Chile : IEEE COMPUTER SOCIETY, Nov. 1997. p.190-198.
66. RICH, E., KNIGHT, K. *Inteligência Artificial*. 2. ed. [s.l.] : McGraw-Hill, 1993.
67. RUSSELL, B. *História do Pensamento Ocidental: a aventura dos Pré-socráticos a Wittgenstein*. Rio de Janeiro : Ediouro, 2001.
68. RUSSELL, Stuart, NORVIG, Peter. *Artificial intelligence: a modern approach*. New Jersey : Prentice Hall, 1995. 932 p.
69. SELLTIZ, C. *et al. Métodos de pesquisa nas relações sociais*. São Paulo : Heder, 1965.
70. SOWA, J. F. Top-layer ontological categories. *International journal of human-computer studies*, 43, p. 669-685, 1995.

71. SOWA, J. F. *Knowledge Representation: Logical, Philosophical and Computational Foundations*. California : Brooks/Cole Publishing Co., 2000.
72. SOWA, J. F. *Building, sharing and merging ontologies*. 2001. Available from World Wide Web: <<http://www.jfsowa.com/ontology/ontoshar.htm>>.
73. SWARTOUT, B., PATIL, R., KNIGHT, K., RUSS, T. Toward Distributed Use of Large-Scale Ontologies. In: WORKSHOP KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS, 20, 1996. Banff, Canada, 1996. *Proceedings...* Banff, Canada, 1996.
74. TAMMA, Valentina, BENCH-CAPON, Trevor. A conceptual model to facilitate knowledge sharing in multi-agent systems. In: WORKSHOP ON ONTOLOGIES IN AGENT SYSTEMS, 29 May 2001, Montreal. *Proceedings of the autonomous agents 2001*, Montreal, 2001. p. 69-76.
75. UNESCO. Guidelines for the establishment and development of monolingual thesauri. Paris, 1973. 37 p.
76. USCHOLD, M. Building Ontologies: Towards a Unified Methodology. proceedings of Expert Systems. In: ANNUAL CONFERENCE OF THE BRITISH COMPUTER SOCIETY SPECIALIST GROUP ON EXPERT SYSTEMS, 16, p. 16-18 Dec. 1996, Cambridge. *Anais...* Cambridge, 1996.
77. USCHOLD, M., GRUNINGER, M. Ontologies: principles, methods and applications. *Knowledge Engineering Review*, v. 11, n. 2, June 1996.
78. VALENTE, A. *Legal knowledge engineering: a modelling approach*. Amsterdam : IOS Press, 1995.
79. VICKERY, B. C. Classificação e indexação nas ciências. Rio de Janeiro : BNG/ Brasilart, 1980. 274 p.
80. WELTY, C., GUARINO, N. Supporting Ontological Analysis of Taxonomic Relationships. *Data and Knowledge Engineering*, v.39, n.1, p. 51-74, 2001.
81. WIEDERHOLD, Gio. Interoperation, Mediation, and Ontologies. In: *INTERNATIONAL SYMPOSIUM ON FIFTH GENERATION COMPUTER SYSTEMS (FGCS)*, Dec. 1994, Tokyo, Japan. *Proceedings Workshop on heterogeneous, cooperative knowledge-bases*. Tokyo, Japan : ICOT, Dec. 1994. v. W3, p. 33-48.

11 ANEXOS

11.1 ANEXO I

Referências bibliográficas da literatura sobre tesouros usadas como fonte material na presente pesquisa

1. ANSI Z39.19:1993. *American National Standard Guidelines for Thesaurus Structure, Construction, and Use*. New York, 1993.
2. CURRAS, Emilia. *Tesouros: linguagens terminológicas*. Brasília : IBICT, 1995. 286 p.
3. HOWERTON, Paul. *Organic and functional concepts of authority files*. In: NEWMAN, Simon M. (Ed). *Information Systems compatibility* . Washington : Spartan Books, 1965.
4. GILCHRIST, Alan. *The thesaurus in retrieval*. London : ASLIB, 1971.
5. UNESCO. *Guidelines for the establishment and development of monolingual thesauri*. Paris, 1973. 37 p.

11.2 ANEXO II

Definições de tesouros, constantes da literatura da área da ciência da informação

1. "[...] vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão (ANSI/NISO Z39-19-1993)."
2. "Tesouro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem – termos, simples ou compostos – encontram-se relacionados entre si sintática e semanticamente (Currás, 1995, p. 88)."
3. "[...] uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou intuitivas (Howerton, 1965, p. 5)".
4. "[...] uma lista autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação coordenada (Alan Gilchrist, 1971, p.6)."
5. "um vocabulário controlado dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento (UNESCO, 1973, p.6)."
6. "[...] No segundo aspecto seria "um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita (UNESCO, 1973, p.6)".

11.3 ANEXO III

Rotação dos termos sobre tesouros obtido pelo sistema kwic

No. Definição	TermoNo. de ocorrências	Termo	texto posterior a ocorrência do termo
	texto anterior a ocorrência do termo		
1	ASSOCIATIVAS.....1 vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e	associativas	entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
3	AUTORIZADA.....2 uma lista	autorizada	, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou intuitivas.
4	uma lista	autorizada	de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação coordenada.
2	COMPOSTOS.....1 Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou	compostos	- encontram-se relacionados entre si sintática e semanticamente.
3	CONCEITO.....1 uma lista autorizada, que pode conduzir o usuário de um	conceito	a outro, por meio de relações heurísticas ou intuitivas.
5	CONHECIMENTO.....1 um vocabulário controlado dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do	conhecimento	.

1	CONTROLADO.....2 vocabulário	controlado	organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
5	um vocabulário	controlado	dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.
6	CONTROLE.....1 um dispositivo de	controle	terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.
4	COORDENADA.....1 uma lista autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação	coordenada	.
5	DINÂMICO.....1 um vocabulário controlado	dinâmico	de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.
6	DOCUMENTAÇÃO.....1 um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de	documentação	, linguagem de informação) mais restrita.
6	DOCUMENTOS.....1 um dispositivo de controle terminológico usado na	documentos	, dos indexadores ou dos usuários numa linguagem do

	tradução da linguagem natural dos		sistema (linguagem de documentação, linguagem de informação) mais restrita.
2	DOCUMENTÁRIOS....1 Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins	documentários	, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
5	DOMÍNIO.....1 um vocabulário controlado dinâmico de termos relacionados semântica e genericamente cobrindo um	domínio	específico do conhecimento.
2	ELEMENTOS.....1 Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os	elementos	lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
1	EQUIVALÊNCIA....1 vocabulário controlado organizado em uma ordem conhecida na qual as relações de	equivalência	, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
2	ESPECIALIZADA....1 Tesauro é uma linguagem	especializada	, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
5	ESPECÍFICO.....1 um vocabulário controlado dinâmico de termos relacionados semântica e genericamente cobrindo um domínio	específico	do conhecimento.
5	GENERICAMENTE....1 um vocabulário controlado dinâmico de termos relacionados semântica e	genericamente	coabrindo um domínio específico do conhecimento.

3	HEURÍSTICAS.....1 uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações	heurísticas	ou intuitivas.
1	HIERÁRQUICAS.....1 vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência,	hierárquicas	e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
1	IDENTIFICADAS....1 vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e	identificadas	por meio de indicadores de relação padrão.
6	INDEXADORES.....1 um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos	indexadores	ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.
4	INDEXAÇÃO.....1 uma lista autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na	indexação	coordenada.
1	INDICADORES.....1 vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de	indicadores	de relação padrão.
6	um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos	indexadores	ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação mais restrita.

6	INFORMAÇÃO.....1 um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de	informação) mais restrita.
3	INTUITIVAS.....1 uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou	intuitivas	.
2	LINGUAGEM.....5 Tesauro é uma	linguagem	especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
6	um dispositivo de controle terminológico usado na tradução da	linguagem	natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.
6	um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa	linguagem	do sistema (linguagem de documentação, linguagem de informação) mais restrita.
6	um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem	de documentação, linguagem de informação) mais restrita.
6	um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação,	linguagem	de informação) mais restrita.

	LINGÜÍSTICOS.....1		
2	Tesouro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos	lingüísticos	que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
	LISTA.....2		
3	uma	lista	autorizada, que pode conduzir o usuário de um conceito a outro, por meio de relações heurísticas ou intuitivas.
4	uma	lista	autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação coordenada.
	LÉXICAS.....1		
4	uma lista autorizada de léxicos, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades	léxicas	, sendo menores, são mais maneáveis e se utilizam na indexação coordenada.
	LÉXICOS.....1		
4	uma lista autorizada de	léxicos	, sem notação, que difere de uma lista de cabeçalhos de assuntos, na qual as unidades léxicas, sendo menores, são mais maneáveis e se utilizam na indexação coordenada.
	NATURAL.....1		
6	um dispositivo de controle terminológico usado na tradução da linguagem	natural	dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.
	NORMALIZADA.....1		
2	Tesouro é uma linguagem especializada,	normalizada	, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.

1	ORDEM.....1 vocabulário controlado organizado em uma	ordem	conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
1	ORGANIZADO.....1 vocabulário controlado	organizado	em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
2	PÓS-COORDENADA...1 Tesouro é uma linguagem especializada, normalizada,	pós-coordenada	, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
2	RELACIONADOS.....2 Tesouro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se	relacionados	entre si sintática e semanticamente.
5	um vocabulário controlado dinâmico de termos	relacionados	semântica e genericamente cobrindo um domínio específico do conhecimento.
1	RELAÇÃO.....1 vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de	relação	padrão.
1	RELAÇÕES.....2 vocabulário controlado organizado em uma ordem	relações	de equivalência, hierárquicas e associativas entre os termos

	conhecida na qual as		são claramente exibidas e identificadas por meio de indicadores de relação padrão. heurísticas ou intuitivas.
3	uma lista autorizada, que pode conduzir o usuário de um conceito a outro, por meio de	relações	
6	RESTRITA.....1 um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais	restrita	.
2	SEMANTICAMENTE...1 Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si sintática e	semanticamente	.
5	SEMÂNTICA.....1 um vocabulário controlado dinâmico de termos relacionados	semântica	e genericamente cobrindo um domínio específico do conhecimento.
2	SIMPLES.....1 Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos,	simples	ou compostos - encontram-se relacionados entre si sintática e semanticamente.
2	SINTÁTICA.....1 Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem - termos, simples ou compostos - encontram-se relacionados entre si	sintática	e semanticamente.
	SISTEMA.....1		

6	um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do	sistema	(linguagem de documentação, linguagem de informação) mais restrita.
6	TERMINOLÓGICO....1 um dispositivo de controle	terminológico	usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.
1	TERMOS.....3 vocabulário controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os	termos	são claramente exibidas e identificadas por meio de indicadores de relação padrão.
2	Tesauro é uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, onde os elementos lingüísticos que compõem -	termos	, simples ou compostos - encontram-se relacionados entre si sintática e semanticamente.
5	um vocabulário controlado dinâmico de	termos	relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.
6	TRADUÇÃO.....1 um dispositivo de controle terminológico usado na	tradução	da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.
3	USUÁRIO.....1 uma lista autorizada, que pode conduzir o	usuário	de um conceito a outro, por meio de relações heurísticas ou intuitivas.
6	USUÁRIOS.....1 um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos,	usuários	numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.

dos indexadores ou dos

1	VOCABULÁRIO.....2	vocabulário	controlado organizado em uma ordem conhecida na qual as relações de equivalência, hierárquicas e associativas entre os termos são claramente exibidas e identificadas por meio de indicadores de relação padrão.
5	um	vocabulário	controlado dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.

11.4 ANEXO IV

Lista de termos encontrados nas definições de tesouros

	TERMOS		OCORRÊNCIA	No. DEF.
relações	ASSOCIATIVAS		1	1
lista	AUTORIZADA		2	3
lista	AUTORIZADA			4
	CONCEITO		1	3
domínio específico do	CONHECIMENTO		1	5
	CONTROLE	terminológico	1	6
indexação	COORDENADA		1	4
vocabulário controlado	DINÂMICO		1	5
linguagem de	DOCUMENTAÇÃO		1	6
linguagem natural dos	DOCUMENTOS		1	6
	DOCUMENTÁRIOS		1	2
	DOMÍNIO	específico do conhecimento	1	5
	ELEMENTOS LINGÜÍSTICOS		1	2
relações de	EQUIVALÊNCIA		1	1
domínio	ESPECÍFICO		1	5
	GENERICAMENTE		1	5
relações	HEURÍSTICAS		1	3
relações	HIERÁRQUICAS		1	1
	INDEXAÇÃO	coordenada	1	4
	INDEXADORES		1	6
	INDICADORES DE RELAÇÃO		1	1
linguagem de	INFORMAÇÃO		1	6
relações	INTUITIVAS		1	3
	LINGUAGEM ESPECIALIZADA		1	2
	LINGUAGEM NATURAL	dos documentos	1	6
	LINGUAGEM	do sistema	1	6
	LINGUAGEM DE DOCUMENTAÇÃO		1	6
	LINGUAGEM	de informação	1	6
	ELEMENTOS LINGÜÍSTICOS		1	2
	LISTA AUTORIZADA		1	3
	LISTA AUTORIZADA	de léxicos	1	4
unidades	LÉXICAS		1	4
	LÉXICOS		1	4
linguagem especializada	NORMALIZADA		1	2
	ORDEM	conhecida	1	1
vocabulário controlado	ORGANIZADO		1	1
linguagem especializada	PÓS-COORDENADA		1	2
termos	RELACIONADOS		1	2
termos	RELACIONADOS			5
	RELAÇÃO	padrão	1	1

	RELAÇÕES	de equivalência, hierárquica e associativas	2	1
	RELAÇÕES	heurísticas e intuitivas		3
linguagem do sistema mais relacionados	RESTRITA		1	6
termos relacionados	SEMANTICAMENTE		1	2
relacionados	SEMÂNTICA		1	5
linguagem do dispositivo de controle	SINTÁTICA		1	2
relações entre os	SISTEMA		1	6
	TERMINOLÓGICO		1	6
	TERMOS		4	1
	TERMOS SIMPLES			2
	TEMOS COMPOSTOS			2
	TERMOS	relacionados semântica		5
	TRADUÇÃO	linguagem natural	1	6
	USUÁRIO		1	3
	USUÁRIOS		1	6
	VOCABULÁRIO CONTROLADO		1	1
	VOCABULÁRIO CONTROLADO			5

11.5 ANEXO V

Lista de termos classificados segundo suas respectivas categorias - tesouros

CATEGORIA "OBJETO"

DOCUMENTOS
DOCUMENTÁRIOS
INFORMAÇÃO

CATEGORIA "LINGUAGEM"

VOCABULÁRIO CONTROLADO [2X]
ELEMENTOS LINGÜÍSTICOS
INDICADORES DE RELAÇÃO
ELEMENTOS LINGÜÍSTICOS
LINGUAGEM ESPECIALIZADA
LINGUAGEM NATURAL
LINGUAGEM [3X]
LINGUAGEM DE DOCUMENTAÇÃO
LINGUAGEM PÓS-COORDENADA
LÉXICAS
LÉXICOS
TERMINOLÓGICO
TERMOS [2X]
TERMOS SIMPLES
TERMOS COMPOSTOS
LISTA AUTORIZADA [2x]
SINTÁTICA

CATEGORIA "ESPAÇO DE CONHECIMENTO"

DOMÍNIO

CATEGORIA "CONTEÚDO SEMÂNTICO"

CONCEITO
CONHECIMENTO
SEMANTICAMENTE
SEMÂNTICA

CATEGORIA "PROCESSOS"

INDEXAÇÃO
TRADUÇÃO

CATEGORIA "ATRIBUTO"

ASSOCIATIVAS
CONTROLE
COORDENADA
DINÂMICO
EQUIVALÊNCIA
ESPECIALIZADA
ESPECÍFICO
GENERICAMENTE
HEURÍSTICAS
HIERÁRQUICAS
INTUITIVAS
NORMALIZADA
ORGANIZADO
RESTRITA

CATEGORIA "SISTEMATIZAÇÃO"

ORDEM
SISTEMA
RELACIONADOS [2X]
RELAÇÃO
RELAÇÕES [2X]

CATEGORIA "AGENTES"

USUÁRIO
USUÁRIOS
INDEXADORES

11.6 ANEXO VI

Referências bibliográficas da literatura sobre ontologias usadas como fonte material na presente pesquisa

1. CHANDRASEKARAN, B., JOSEPHSON, J. R., BENJAMINS, V. R.. What are Ontologies, and why do we need them. *IEEE Intelligent Systems*, v.14, n. 1, p.20-26, Jan/Feb 1999.
2. ERIKSSON, Henrik, PUERTA, Angel R., MUSEN, Mark A. Generation of knowledgeacquisition tools from domain ontologies. In: SYSTEMS WORKSHOP, BANFF KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED, 8, 1994, Banff, Canada. v. 1, p. 7-20, Banff, Canada, 1994.
3. GRUBER, T. R. *Ontolingua: a mechanism to support portable ontologies*. Stanford : *Knowledge Systems Laboratory* Stanford University, 1992. (Relatório KSL91-66).
4. GRUNINGER, M, FOX, M. S. Methodology for the design and evaluation of ontologies. In INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1995, *Montreal, Quebec, Canada*. *Proceedings... Montreal, Quebec, Canada*, 1995.
5. GUARINO, Nicola, GIARETTA, P. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: N. MARS (Ed.) *Towards very large Knowledge Bases: Knowledge Building and Knowledge sharing*, 1995. Amsterdam : IOS Press. p. 25-32.
6. GUARINO, N. Formal Ontology, Conceptual Analysis and Knowledge Representation. In: GUARINO, Nicola, POLI, Roberto (Eds).. *International Journal of Human-Computer Studies*, v. 43, n. 5/6, 1995b. Número especial.
7. HOVY, E. Combining and standardizing large-scale, practical ontologies for machine translation and other uses. In: LANGUAGE RESOURCES AND EVALUATION CONFERENCE. Granada, Spain, May 1998. *Conferências... Granada, Spain, May 1998*.
8. MENA, E., KASHYAP, V., SHETH, A. *et al.* OBSERVER: an approach for query processing in Global Information systems based on Interoperation across pre-existing Ontologies. In: IFCIS - INTERNATIONAL CONFERENCE ON COOPERATIVE INFORMATION SYSTEMS (COOPIS '96), 1996. *Proceedings...*, Brussels, Belgium, June 1996.

9. SWARTOUT, B., PATIL, R., KNIGHT, K., RUSS, T. Toward Distributed Use of Large-Scale Ontologies. In: WORKSHOP KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS, 20, 1996. Banff,Canada, 1996. *Proceedings...* Banff,Canada, 1996.
- 10.USCHOLD, M. Building Ontologies: Towards a Unified Methodology. proceedings of Expert Systems. In: ANNUAL CONFERENCE OF THE BRITISH COMPUTER SOCIETY SPECIALIST GROUP ON EXPERT SYSTEMS, 16, p. 16-18 December, 1996, Cambridge. *Anais...*, Cambridge, 1996.
- 11.USCHOLD, M., GRUNINGER, M.. Ontologies: principles, methods and applications. *Knowledge Engineering Review*, v. 11, n. 2, 1996.
- 12.WIEDERHOLD, Gio. Interoperation, Mediation, and Ontologies. In: *INTERNATIONAL SYMPOSIUM ON FIFTH GENERATION COMPUTER SYSTEMS (FGCS)*, Dec. 1994, Tokyo, Japan. *Proceedings Workshop on heterogeneous, cooperative knowledge-bases*. Tokyo, Japan : ICOT, Dec. 1994. v. W3, p. 33-48.

11.7 ANEXO VII

Definições de ontologias, constantes da literatura da área da ciência da computação

1. "[...] Ontologia é um vocabulário de representação, frequentemente especializado para algum domínio ou assunto." (CHANDRASEKARAN *et al.*, 1999, p. 1)
2. "[...] Ontologia é usada para referir a um corpo de conhecimento descrevendo algum domínio, tipicamente um conhecimento comum de um domínio, usando um vocabulário de representação." (CHANDRASEKARAN *et al.*, 1999, p. 1)
3. "Uma Ontologia é um modelo declarativo dos termos e relações em um domínio." (ERIKSSON *et al.*, 1994, p.1)
4. "[...] Ontologia como um sistema conceitual relacionado a uma base de conhecimento." (GUARINO e GIARETTA, 1995. p. 1).
5. "Ontologia como uma representação de um sistema conceitual". (GUARINO e GIARETTA, 1995, p. 1).
6. É uma "especificação formal e explícita de uma conceitualização." (GRUBER, 1993, p.1)
7. "[...] Uma ontologia é uma descrição formal de entidades e suas propriedades, relações, restrições e comportamentos." (GRUNINGER e FOX, 1995, p.1)
8. "Uma ontologia é um conjunto de termos, associado com definições em linguagem natural e, se possível, usando relações formais e restrições, sobre um domínio de interesse." (HOVY, 1998, p.2)
9. "Cada ontologia é um conjunto de termos de interesse em um domínio particular, expressos em lógica de descrições." (MENA *et al.*, 1996, p.3)
10. "[...] Uma ontologia é um conjunto de termos estruturados hierarquicamente para descrever um domínio que pode ser usado como um esquema base para uma base de conhecimento." (SWARTOUT *et al.*, 1996, p.1)
11. "Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma

especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos." (USCHOLD, 1996, p. 3)

12. "Ontologias são acordos sobre conceitualizações compartilhadas." (USCHOLD e GRUNNINGER, 1996, p.6)

13. "[...] uma ontologia é um vocabulário de termos e uma especificação de suas relações." (WIEDERHOLD, 1994, p. 6)

11.8 ANEXO VIII

Rotação dos termos sobre ontologias obtido pelo sistema kwic

TermoNo. de ocorrências		Termo	
No. Definição	texto anterior a ocorrência do termo		texto posterior a ocorrência do termo
ASSUNTO.....1 1	Ontologia é um vocabulário de representação, frequentemente especializado para algum domínio ou	assunto	.
BASE.....3			
4	Ontologia como um sistema conceitual relacionado a uma	base	de conhecimento
10	Uma ontologia é um conjunto de termos estruturados hierarquicamente para descrever um domínio que pode ser usado como um esquema base para uma	base	de conhecimento.
COMPORTAMENTOS..1 7	Uma ontologia é uma descrição formal de entidades e suas propriedades, relações, restrições e	comportamentos	
CONCEITOS.....1			
11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui	conceitos	estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos

	definições e uma indicação de como		
CONCEITUAL.....2			
4	Ontologia como um sistema	conceitual	relacionado a uma base de conhecimento
5	Ontologia como uma representação de um sistema	conceitual	
CONCEITUALIZAÇÃO.....1			
6	especificação formal e explícita de uma	conceitualização	
CONCEITUALIZAÇÕES....1			
12	Ontologias são acordos sobre	conceitualizações	compartilhadas.
CONHECIMENTO.....4			
4	Ontologia como um sistema conceitual relacionado a uma base de	conhecimento	
10	Uma ontologia é um conjunto de termos estruturados hierarquicamente para descrever um domínio que pode ser usado como um esquema base para uma base de	conhecimento	.
2	Ontologia é usada para referir a um corpo de	conhecimento	descrevendo algum domínio, tipicamente um conhecimento sendo comum de um domínio, usando um vocabulário de representação.
2	Ontologia é usada para referir a um corpo de conhecimento descrevendo algum domínio, tipicamente um	conhecimento	sendo comum de um domínio, usando um vocabulário de representação.
DECLARATIVO.....1			

3	Uma Ontologia é um modelo	declarativo	. dos termos e relações em um domínio
DEFINIÇÕES.....2			
11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui	definições	e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos
8	Uma ontologia é um conjunto de termos, associado com	definições	em linguagem natural e, se possível, usando relações formais e restrições, sobre um domínio de interesse.
DESCREVENDO.....1			
2	Ontologia é usada para referir a um corpo de conhecimento	descrevendo	algum domínio, tipicamente um conhecimento sendo comum de um domínio, usando um vocabulário de representação.
DESCREVER.....1			
10	Uma ontologia é um conjunto de termos estruturados hierarquicamente para	descrever	um domínio que pode ser usado como um esquema base para uma base de conhecimento.
DESCRIÇÃO.....1			
7	Uma ontologia é uma	descrição	formal de entidades e suas propriedades, relações, restrições e comportamentos.
DESCRIÇÕES.....1			
9	Cada ontologia é um conjunto de termos de interesse em um domínio particular, expressos em lógica de	descrições	
DOMÍNIO.....8			
11	Uma ontologia pode possuir uma variedade de formas,	domínio	e restringe as possíveis interpretações dos

		mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no		termos
9		Cada ontologia é um conjunto de termos de interesse em um	domínio	particular, expressos em lógica de descrições
10		Uma ontologia é um conjunto de termos estruturados hierarquicamente para descrever um	domínio	que pode ser usado como um esquema base para uma base de conhecimento.
1		Ontologia é um vocabulário de representação, frequentemente especializado para algum	domínio	ou assunto.
2		Ontologia é usada para referir a um corpo de conhecimento descrevendo algum	domínio	, tipicamente um conhecimento sendo comum de um domínio, usando um vocabulário de representação.
2		Ontologia é usada para referir a um corpo de conhecimento descrevendo algum domínio, tipicamente um conhecimento sendo comum de um	domínio	, usando um vocabulário de representação.
8		Uma ontologia é um conjunto de termos, associado com definições em linguagem natural e, se possível, usando relações formais e restrições, sobre um	domínio	de interesse.
3		Uma Ontologia é um modelo declarativo dos termos e relações em um	domínio	.
ENTIDADES.....1 7		Uma ontologia é uma descrição formal de	entidades	e suas propriedades, relações, restrições e comportamentos.

ESPECIALIZADO.....1 1	Ontologia é um vocabulário de representação, frequentemente	especializado	para algum domínio ou assunto.
ESPECIFICAÇÃO.....3 11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma	especificação	de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos
6		especificação	formal e explícita de uma conceitualização
13	Uma ontologia é um vocabulário de termos e uma	especificação	de suas relações.
ESQUEMA.....1			
10	Uma ontologia é um conjunto de termos estruturados hierarquicamente para descrever um domínio que pode ser usado como um	esquema	base para uma base de conhecimento.
ESTRUTURA.....1 11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma	estrutura	no domínio e restringe as possíveis interpretações dos termos
ESTRUTURADOS.....1 10	Uma ontologia é um conjunto de termos	estruturados	hierarquicamente para descrever um

			domínio que pode ser usado como um esquema base para uma base de conhecimento.
EXPLÍCITA.....1			
6	especificação formal e	explícita	de uma conceitualização.
EXPRESSOS.....1			
9	Cada ontologia é um conjunto de termos de interesse em um domínio particular,	expressos	em lógica de descrições
FORMAIS.....1			
8	Uma ontologia é um conjunto de termos, associado com definições em linguagem natural e, se possível, usando relações	formais	e restrições, sobre um domínio de interesse.
FORMAL.....2			
6	especificação	formal	e explícita de uma conceitualização
7	Uma ontologia é uma descrição	formal	de entidades e suas propriedades, relações, restrições e comportamentos.
HIERARQUICAMENTE.....1			
10	Uma ontologia é um conjunto de termos estruturados	hierarquicamente	para descrever um domínio que pode ser usado como um esquema base para uma base de conhecimento.
INTER-RELACIONADOS...1			
11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão	inter-relacionados	o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos
INTERPRETAÇÕES.....1			
11	Uma ontologia pode possuir uma	interpretações	dos termos

variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis

LINGUAGEM.....1 8	Uma ontologia é um conjunto de termos, associado com definições em	linguagem	natural e, se possível, usando relações formais e restrições, sobre um domínio de interesse.
LÓGICA.....1 9	Cada ontologia é um conjunto de termos de interesse em um domínio particular, expressos em	lógica	de descrições
MODELO.....1 3	Uma Ontologia é um	modelo	declarativo dos termos e relações em um domínio.
NATURAL.....1 8	Uma ontologia é um conjunto de termos, associado com definições em linguagem	natural	e, se possível, usando relações formais e restrições, sobre um domínio de interesse.
PROPRIEDADES.....1 7	Uma ontologia é uma descrição formal de entidades e suas	propriedades	, relações, restrições e comportamentos.
RELAÇÕES.....4 7	Uma ontologia é uma descrição formal de entidades e suas propriedades,	relações	, restrições e comportamentos.
13	Uma ontologia é um vocabulário de termos e uma especificação de suas	relações	.
8	Uma ontologia é um	relações	formais e restrições,

	conjunto de termos, associado com definições em linguagem natural e, se possível, usando		sobre um domínio de interesse.
3	Uma Ontologia é um modelo declarativo dos termos e	relações	em um domínio.
REPRESENTAÇÃO.....3			
5	Ontologia como uma	representação	de um sistema conceitual
1	Ontologia é um vocabulário de	representação	, frequentemente especializado para algum domínio ou assunto.
2	Ontologia é usada para referir a um corpo de conhecimento descrevendo algum domínio, tipicamente um conhecimento sendo comum de um domínio, usando um vocabulário de	representação	
RESTRINGE.....1			
11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e	restringe	as possíveis interpretações dos termos
RESTRITÕES.....2			
7	Uma ontologia é uma descrição formal de entidades e suas propriedades, relações,	restritões	e comportamentos.
8	Uma ontologia é um conjunto de termos, associado com definições em linguagem natural e, se possível, usando relações formais e	restritões	, sobre um domínio de interesse.

SIGNIFICADOS.....1 11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus	significados	. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos
SISTEMA.....2 4	Ontologia como um	sistema	conceitual relacionado a uma base de conhecimento conceitual
5	Ontologia como uma representação de um	sistema	
TERMOS.....7 11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de	termos	, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos
11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um vocabulário de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos	termos	
9	Cada ontologia é um conjunto de	termos	de interesse em um domínio particular, expressos em lógica de descrições estruturados
10	Uma ontologia é um	termos	

	conjunto de		hierarquicamente para descrever um domínio que pode ser usado como um esquema base para uma base de conhecimento.
13	Uma ontologia é um vocabulário de	termos	e uma especificação de suas relações.
8	Uma ontologia é um conjunto de	termos	, associado com definições em linguagem natural e, se possível, usando relações formais e restrições, sobre um domínio de interesse.
3	Uma Ontologia é um modelo declarativo dos	termos	e relações em um domínio.
VOCABULÁRIO.....4			
11	Uma ontologia pode possuir uma variedade de formas, mas necessariamente incluirá um	vocabulário	de termos, e alguma especificação de seus significados. Isto inclui definições e uma indicação de como conceitos estão inter-relacionados o que impõem uma estrutura no domínio e restringe as possíveis interpretações dos termos
1	Ontologia é um	vocabulário	de representação, frequentemente especializado para algum domínio ou assunto.
2	Ontologia é usada para referir a um corpo de conhecimento descrevendo algum domínio, tipicamente um conhecimento sendo comum de um domínio, usando um	vocabulário	de representação.
13	Uma ontologia é um	vocabulário	de termos e uma especificação de suas relações.

11.9 ANEXO IX

Lista de termos encontrados nas definições de ontologias

	TERMOS		Ocorrência	NO. DEF.
domínio ou relacionado	ASSUNTO		1	1
	BASE DE CONHECIMENTO		2	4
	BASE DE CONHECIMENTO			10
conceitualizações	COMPARTILHADAS		1	12
	COMPORTAMENTOS		1	7
	CONCEITOS		1	11
sistema	CONCEITUAL		2	4
sistema	CONCEITUAL			5
	CONCEITUALIZAÇÃO	compartilhada	1	6
	CONCEITUALIZAÇÕES	compartilhadas	1	12
	CONHECIMENTO		1	2
corpo	CONHECIMENTO		1	2
	DECLARATIVO		1	3
	DEFINIÇÕES		2	8
	DEFINIÇÕES			11
	DESCREVENDO	algum domínio	1	2
	DESCREVER	um domínio	1	10
	DESCRIÇÃO	formal de entidades	1	7
estrutura no	DOMÍNIO		8	11
	DOMÍNIO	particular		9
descrever um	DOMÍNIO			10
	DOMÍNIO	ou assunto		1
	DOMÍNIO			2
	DOMÍNIO			2
	DOMÍNIO	de interesse		8
	DOMÍNIO			3
relações em um	DOMÍNIO			
descrição formal de	ENTIDADES		1	7
	ESPECIALIZADO		1	1
	ESPECIFICAÇÃO	de significados	3	11
	ESPECIFICAÇÃO	formal de uma conceitualização		6
	ESPECIFICAÇÃO	de suas relações		13
	ESQUEMA	base	1	10
	ESTRUTURA	no domínio	1	11
termos	ESTRUTURADOS		1	10
especificação formal e	EXPLÍCITA	de uma conceitualização	1	6
	EXPRESSOS	em lógica de descrições	1	9
relações	FORMAIS		1	8
especificação	FORMAL		2	6
descrição	FORMAL			7
termos estruturados	HIERARQUICAMENTE		1	10
conceitos	INTER-RELACIONADOS		1	11
	INTERPRETAÇÕES	dos termos	1	11
	LINGUAGEM natural		1	8

	LÓGICA DE DESCRIÇÕES		1	9
	MODELO	declarativo dos termos	1	3
	PROPRIEDADES		1	7
	RELAÇÕES		4	7
especificação de suas	RELAÇÕES			13
	RELAÇÕES	formais		8
	RELAÇÕES	em um domínio		3
	REPRESENTAÇÃO	de um sistema conceitual	3	5
vocabulário de	REPRESENTAÇÃO			1
vocabulário de	REPRESENTAÇÃO			2
	RESTRINGE	interpretações dos termos	1	11
	RESTRICÇÕES		2	7
	RESTRICÇÕES			8
	SIGNIFICADOS		1	11
	SISTEMA	conceitual	2	4
	SISTEMA	conceitual		5
vocabulário de	TERMOS		7	11
vocabulário de	TERMOS			13
interpretações dos	TERMOS			11
conjunto de	TERMOS			9
conjunto de	TERMOS			10
conjunto de	TERMOS			8
modelo declarativo	TERMOS			3
	VOCABULÁRIO	de termos	4	11
	VOCABULÁRIO	de representação		1
	VOCABULÁRIO	de representação		2
	VOCABULÁRIO	de termos		13

11.10ANEXO X

Lista de termos classificados segundo suas respectivas categorias - ontologias

CATEGORIA "OBJETO"

BASE DE CONHECIMENTO [2x]
ENTIDADES

CATEGORIA "LINGUAGEM"

LÓGICA DE DESCRIÇÕES
LINGUAGEM NATURAL
TERMOS [7X]
VOCABULÁRIO [4X]

CATEGORIA "ESPAÇO DE CONHECIMENTO"

ASSUNTO
DOMÍNIO [8X]
CORPO DE CONHECIMENTO

CATEGORIA "CONTEÚDO SEMÂNTICO"

CONCEITOS
CONCEITUALIZAÇÃO / CONCEITUALIZAÇÕES
DEFINIÇÕES [2X]
CONHECIMENTO
SIGNIFICADOS

CATEGORIA "PROCESSOS"

DESCREVENDO
DESCREVER
DESCRIÇÃO
ESPECIFICAÇÃO [3X]
REPRESENTAÇÃO [3X]
INTERPRETAÇÕES

CATEGORIA "ATRIBUTO"

COMPARTILHADAS
DECLARATIVO
HIERARQUICAMENTE
FORMAIS
FORMAL [2X]
EXPLÍCITA
EXPRESSOS
ESPECIALIZADO
PROPRIEDADES
COMPORTAMNETOS
RESTRINGE
RESTRICÇÕES [2X]

CATEGORIA "SISTEMATIZAÇÃO"

ESQUEMA
ESTRUTURA
ESTRUTURADOS
MODELO
RELAÇÕES [4X]
SISTEMA CONCEITUAL [2X]
INTER-RELACIONADOS