

**INSTITUTO BRASILEIRO DE CIÊNCIA E TECNOLOGIA
UNIVERSIDADE FEDERAL FLUMINENSE**

LUANA FARIAS SALES

**ONTOLOGIAS DE DOMÍNIO
ESTUDO DAS RELAÇÕES CONCEITUAIS E SUA APLICAÇÃO**

**Niterói
2006**

LUANA FARIAS SALES

ONTOLOGIAS DE DOMÍNIO
ESTUDO DAS RELAÇÕES CONCEITUAIS E SUA APLICAÇÃO

Dissertação apresentada ao Curso de Pós-Graduação em Ciência da Informação do convênio IBICT – UFF, como pré-requisito parcial para obtenção do Grau de Mestre. Área de Concentração: O conhecimento da informação e a informação para o conhecimento; Linha de Pesquisa: Representação, gestão e tecnologia da informação.

ORIENTADOR: PROF. DR. MARIA LUIZA DE ALMEIDA CAMPOS
CO-ORIENTADORA: PROF. DR. HAGAR ESPANHA GOMES

Niterói
2006

S163 Sales, Luana Farias.

Ontologias de domínio: estudo das relações conceituais e sua aplicação / Luana Farias Sales – Niterói: UFF, 2006.

141f.

Dissertação (Mestrado em Ciência da Informação) –
Universidade Federal Fluminense, 2006.

1. Ontologias. 2. Relações conceituais. 3. Recuperação de Informação. I. Título.

CDD
CDU

LUANA FARIAS SALES

ONTOLOGIAS DE DOMÍNIO
ESTUDO DAS RELAÇÕES CONCEITUAIS E SUA APLICAÇÃO

Dissertação apresentada ao Curso de Pós-Graduação em Ciência da Informação do convênio IBICT – UFF, como pré-requisito parcial para obtenção do Grau de Mestre. Área de Concentração: O conhecimento da informação e a informação para o conhecimento; Linha de Pesquisa: Representação, gestão e tecnologia da informação.

Data de Aprovação: _____ \ _____ \ _____

Prof. Maria Luiza de Almeida Campos - Orientadora
D.Sc. Ciência da Informação – UFRJ/IBICT

Data de Aprovação: _____ \ _____ \ _____

Prof. Hagar Espanha Gomes – Co-orientadora
Livre Docente – UFF

Data de Aprovação: _____ \ _____ \ _____

Prof. Carlos Henrique Marcondes
D.Sc. Ciência da Informação – UFRJ/IBICT

Data de Aprovação: _____ \ _____ \ _____

Prof. Maria Luiza Machado Campos
Ph.D. Information Systems, University of East Anglia, Inglaterra

Data de Aprovação: _____ \ _____ \ _____

Prof. Rosa Inês de Novais Cordeiro
Doutora em Comunicação e Cultura - UFRJ/Eco
Pós-Doutorado no Instituto de Psicologia da UFRJ

Às mestras Maria Luiza e Hagar, por me ensinarem o caminho da pesquisa com disciplina e responsabilidade.

À todas as pessoas que de alguma forma incentivaram a realização desse projeto.

Agradecimentos

Agradeço primeiramente a Deus por ter me conduzido a esse caminho.

Agradeço às minhas orientadoras Maria Luiza de Almeida Campos e Hagar Espanha Gomes, por me acompanharem nesse caminho.

Agradeço aos professores do PPGCI, em especial aos professores da UFF - Carlos Henrique Marcondes, Sandra Rebel, Lídia Freitas, Rosa Inês de Novais - que vêm me acompanhando desde a Graduação.

À grande amiga Dilza Motta, pelas dicas sempre tão pertinentes e principalmente pelas palavras de ânimo e de carinho.

A todos os colegas de curso da turma de mestrado de 2004 e 2005, em especial, Dilza Bastos, Lúcia Maria, Ingrid Beck, Roberta Pereira da Silva, Laffayete Alvares Junior, Joyce Cardoso, Ildenise Novo e Luciana Gracioso.

A todos os funcionários do IBICT e UFF, em especial, à bibliotecária Sônia Burnier, pelo auxílio na busca de informação.

Ao amigo Rogério Ramalho pelas trocas virtuais.

Aos amigos bioinformatas: Pablo Mendes, Linair Campos, Ana Carolina, Glauber e Diogo Mattos, já que não consigo mais separá-los em biólogos e informatas.

À professora Maria Luiza Machado Campos, por acreditar neste projeto e ter aberto espaços para discussões interdisciplinares.

Às amigas Simone Rabello e Heloísa Helena, por me incentivarem a trilhar esse caminho, desde a graduação.

Aos novos amigos Almir Barbio, Maria Bernarda e Elaine Rosa, pela compreensão nessa reta final.

Aos meus pais, Ulisses e Dalva, minha irmã Anna Lú, por me acompanharem em todas essas conquistas.

À minha “prima-irmã” Anna Maura, pelos momentos de desconcentração e minha “mãedrinha” por, mais uma vez, me receber e cuidar de mim.

Ao Jorge Rogério, meu amor, que mesmo nos momentos distantes, me incentivou para que eu não desistisse dos meus sonhos.

“Tudo está em relação com tudo. Nada está isolado, existindo solitário, de si para si. Tudo co-existe e inter-existe com todos os outros seres do universo”
(BOFF, 1997, p. 72)

“Há efetivamente, necessidade de um pensamento:

- que compreenda que o conhecimento das partes depende do conhecimento do todo e que o conhecimento do todo depende do conhecimento das partes;
- que reconheça e examine os fenômenos multidimensionais, em vez de isolar, de maneira mutiladora, cada uma de suas dimensões.;
- que reconheça e trate as realidades, que são, concomitantemente solidárias e conflituosas;
 - que respeite a diferença, enquanto reconhece a unicidade”

(MORIN, 2003, p. 89)

RESUMO

No âmbito da elaboração de Ontologias, uma problemática que tem se colocado é a ausência de um padrão teórico-metodológico para sua elaboração, o que implica na criação de regras de inferências consistentes que possam ser interpretáveis pela máquina. Propõe-se então uma sistematização das relações apresentadas nas literaturas da Ciência da Informação, da Terminologia, da Ciência da Computação e da Bioinformática. Na Ciência da Informação, as Categorias Fundamentais proporcionam um modelo de representação de um domínio, mas não explicitam a relação entre elas. A Ciência da Computação explicita as relações, mas não fornece um contexto. A sistematização realizada busca unir as duplas de categorias (relações categoriais) com as relações potenciais entre elas (relações formais). A partir da seleção de um corpus da Gene Ontology, foi feita uma análise das definições, a fim de identificar as relações já mencionadas na literatura ou descobrir a existência de novas relações. Apesar da Gene Ontology apresentar padrões para as definições, nem sempre foi possível identificar facilmente as relações. Conclui-se que relações diferentes podem surgir em diferentes domínios e que definições sistematizadas são indispensáveis para o estabelecimento seguro de relações formais.

Descritores: Ontologias. Relações conceituais. Recuperação de informação.

ABSTRACT

Lack of theoretical bases and sound methodologies are problems that arise when building Ontologies. These constitute hindrance to the creation of inference rules for machine understanding. The study proposes systematization of relations found in the literature from Information Science, Terminology, Computer Science and Bioinformatics. Fundamental Categories in Information Science provide a domain representation model, but not the relation among them. In Computer Science relations are revealed, but not the context in which they occur. The systematization proposed aims at linking couple of categories (categorical relations) with relations properly (formal relations). A sample of definitions from Gene Ontology was analyzed to identify relations already mentioned in the literature and/or new ones. Despite the models for definition in Gene Ontology not always relations could be easily identified. It is concluded that different relations may arise in different domains and that systematic definitions are fundamental for the establishment of formal relations.

Descriptors: Ontologies. Conceptual relations. Information retrieval.

LISTA DE FIGURAS

Figura 1 – Tipos de Ontologias.....	30
Figura 2 – Planilhas do banco de dados Interpro.....	44

LISTA DE QUADROS

Quadro 1- Relacionamentos Não-hierárquicos apresentados por Neelameghan..	49
Quadro 2- Classificação das Relações segundo a TGT.....	56
Quadro 3- Relações Complexas formais.....	60
Quadro 4- Relações Complexas categoriais.....	61
Quadro 5- Categorias do UMLS (parte 1).....	80
Quadro 6- Categorias do UMLS (parte 2).....	81
Quadro 7- Relações do UMLS.....	82
Quadro 8- Sistematização das Relações Genéricas.....	84
Quadro 9- Sistematização das Relações Partitivas.....	85
Quadro 10- Sistematização das Relações Funcionais.....	88
Quadro 11- Padrão de Definição da Categoria Função Molecular.....	94
Quadro 12- Padrão de Definição da Categoria Processo Biológico	95
Quadro 13- Uso de Qualificadores na GO	96
Quadro 14- Padrão de Definição da Categoria Componente Celular.....	98
Quadro 15- Sistematização das Relações da GO.....	100

SUMÁRIO

1	CONSIDERAÇÕES INICIAIS: AS ONTOLOGIAS NO CONTEXTO DA WEB.....	13
2	OBJETIVOS.....	18
2.1	OBJETIVO GERAL.....	18
2.2	OBJETIVOS ESPECÍFICOS.....	18
3	ONTOLOGIA.....	19
3.1	ORIGEM DO TERMO, DEFINIÇÃO E COMPOSIÇÃO DAS ONTOLOGIAS.....	22
3.2	POSSIBILIDADES DE APLICAÇÃO DAS ONTOLOGIAS.....	26
3.3	TIPOS DE ONTOLOGIAS E SUAS APLICAÇÕES.....	29
3.4	METODOLOGIAS PARA CONSTRUÇÃO DE ONTOLOGIAS.....	32
4	A GENE ONTOLOGY: UMA ONTOLOGIA DE DOMÍNIO.....	40
5	RELAÇÕES CONCEITUAIS NO ÂMBITO DA CIÊNCIA DA INFORMAÇÃO, DA TERMINOLOGIA E DA CIÊNCIA DA COMPUTAÇÃO.....	46
5.1	RELAÇÕES CONCEITUAIS NA CIÊNCIA DA INFORMAÇÃO.....	47
5.2	RELAÇÕES CONCEITUAIS NA TERMINOLOGIA.....	54
5.3	RELAÇÕES CONCEITUAIS NA CIÊNCIA DA COMPUTAÇÃO.....	62
5.4	COMPARAÇÃO ENTRE AS TEORIAS PARA O ESTABELECIMENTO DE RELAÇÕES.....	70
6	RELAÇÕES CONCEITUAIS NA ÁREA BIOMÉDICA.....	75
7	SISTEMATIZAÇÃO DAS RELAÇÕES.....	83
7.1	RELAÇÕES GENÉRICAS.....	84
7.2	RELAÇÕES PARTITIVAS	84
7.3	RELAÇÕES FUNCIONAIS.....	86
8	APLICAÇÃO DO MODELO DE RELAÇÕES NO ÂMBITO DA GENE ONTOLOGY.....	91
8.1	A AMOSTRA SELECIONADA	92
8.2	RELAÇÕES CATEGORIAIS.....	93
8.3	RELAÇÕES FORMAIS.....	93
8.3.1	Função molecular.....	94
8.3.2	Processo biológico.....	95
8.3.3	Componente celular.....	98
8.4	SISTEMATIZAÇÃO DAS RELAÇÕES.....	99
9	CONSIDERAÇÕES FINAIS.....	102
	REFERÊNCIAS.....	105

APÊNDICE A.....	116
APÊNDICE B.....	126
APÊNDICE C.....	130
ANEXO A.....	139

1 CONSIDERAÇÕES INICIAIS: AS ONTOLOGIAS NO CONTEXTO WEB

Quando a World Wide Web foi criada, seu propósito inicial era prover acesso à informação pela comunidade científica, através de um projeto global de hipertexto, designado para permitir que as pessoas trabalhassem juntas por meio de organizações, links e navegações de páginas de conteúdo.

Esse projeto foi criado por Tim Berners-Lee, enquanto trabalhava no laboratório europeu de Física CERN (European Organization for Nuclear Research Center), em 1989 e foi chamado de "World Wide Web" (CAILLIAU; GILLIES, 2006; RAMALHO, 2006). Porém, um espaço que foi criado inicialmente para troca de informação científica tomou uma proporção muito maior, devido a características como "liberdade de publicação, autonomia das fontes e controle descentralizado", que proporcionaram uma grande variedade de páginas publicadas.

Essa diversidade de páginas mudou o comportamento dos usuários, ampliando assim as formas de utilização desse instrumento. Atualmente, há uma "significativa dependência dos serviços prestados via este novo ambiente de interação", tendo em vista que usuários agora utilizam a rede para busca de informação de todo tipo, por exemplo, consulta bancária, marcação de vôos, consulta sobre andamento de processos etc. Com isso, a Web passa a ser então uma necessidade na vida do cidadão (CAMPOS; CAMPOS; CAMPOS, 2006, p. 55).

Uma questão porém vem perpassando todas as fases da Web: a recuperação precisa de informações. Quanto maior o número de informações publicadas, maior a dificuldade em acessá-las e quanto maior a utilização desse instrumento, maior é a necessidade sentida de se obter a informação desejada em menor espaço de tempo.

Esta problemática é ocasionada, entre outros fatores, pela ausência de padrões para a publicação de recursos na Web e devido a isto, o criador da Web atual, Tim Berners-Lee, se juntou a outros pesquisadores em 1994 para criação de um Consórcio Internacional chamado W3C¹ - World Wide Web Consortium.

A este consórcio se unem empresas, instituições acadêmicas, profissionais e cientistas, com o objetivo de desenvolver novos padrões para publicação na Web e também novas tecnologias que possibilitem não apenas o processamento de informação por pessoas, mas também por software.

A Web Semântica seria, então, uma nova geração da Web atual e segundo seus idealizadores, “ela visa fornecer estruturas e dar significado semântico ao conteúdo das páginas Web, criando um ambiente onde agentes de software e usuários possam trabalhar de forma cooperativa” (BERNERS-LEE; HENDLER; LASSILA, 2001). Neste sentido, a Web poderá auxiliar o usuário a acessar a informação que deseja, podendo algumas vezes ajudá-lo a decidir sobre qual serviço utilizar.

A Web Semântica vem, então, sugerir que a integração e a interoperabilidade dos dados sejam feitas em camadas. Esta visão de camadas, conforme Ramalho (2006), está em constante mutação, mas atualmente pode-se falar em cinco camadas específicas: Camada Estrutural; Camada Sintática; Camada Semântica; Camada Lógica e Camada de Confiança.

Moura ([2002], p. 2), no entanto, afirma, segundo o W3C, que a interoperabilidade dos dados deve acontecer nos três primeiros níveis: estrutural, sintático e semântico. Segundo esta autora, a trilogia de níveis de interoperabilidade pode ser definida da seguinte forma: o **nível sintático** é aquele que “determina como os metadados devem ser codificados para a transferência de informações”; o **nível estrutural** é aquele que “especifica como os recursos estão organizados, juntamente com os tipos de recursos envolvidos e os possíveis valores para cada tipo”. E, finalmente, o **nível semântico**, que é “aquele que possibilita a compreensão de cada elemento descritor do recurso com as associações nele embutidas”, através do uso de vocabulários específicos.

Para atender o nível sintático, o consórcio W3C, sugere o uso do XML, em complemento à linguagem html, pois enquanto esta é voltada para o designer da página, aquela é voltada para o intercâmbio de dados e a comunicação entre

¹ Cf. <<http://www.w3.org/>>.

sistemas. Outra característica do XML é que este descreve o próprio conteúdo do documento, e, além disso permite que o usuário defina as suas próprias tags para criar uma estrutura.

Se, por um lado, permitir que o usuário criasse suas tags fez do XML uma linguagem mais flexível, por outro, essa flexibilidade permitiu a geração de algumas ambigüidades semânticas, já que não existe um padrão de tags para a descrição dos elementos. De qualquer forma, o XML foi criado para proporcionar a interoperabilidade sintática e por isso o Consórcio W3C sugere que a interoperação seja realizada em três níveis distintos, a fim de que eles possam se complementar.

Para a interoperação no nível estrutural, o consórcio sugere uma linguagem chamada RDF, ou seja, o Resource Description Framework, que é uma linguagem para a especificação de relacionamentos entre dados, ou mais precisamente, o RDFS, que é um esquema RDF, que representa as relações entre os dados, através de um DLG, isto é, Directed Labeled Graphs.

Para atender o nível semântico, eles retomam uma proposta da Inteligência Artificial, que é a criação dos vocabulários controlados, com linguagem própria de máquina, denominados **Ontologias** (W3C, 2001).

As Ontologias possuem uma parte terminológica, composta de termos, definições e relações, e também uma parte processável por máquina, expressa em linguagem formal, com regras de inferências, relacionamentos e definições expressas nessa linguagem.

A questão que impulsionou a elaboração do projeto de Mestrado que veio a dar origem a este estudo era: como a Ciência da Informação poderia contribuir para elaboração de Ontologias?

A inexistência de uma metodologia unificada para construção de ontologias foi apontada como problema por diversos autores (FERNANDEZ-LÓPEZ, 1999; JONES; BENCH-CAPON; VISSER, 1998). Entre os problemas ocasionados pela falta de metodologia, poderíamos citar, entre outros, ausência de um padrão para construção de definições, bem como de um padrão para o estabelecimento de relações.

O presente estudo aborda especificamente a questão da determinação de relações consistentes no âmbito da elaboração de ontologias.

São as relações entre conceitos, estabelecidas de forma coerente, que vão permitir a criação de regras de inferência consistentes, possibilitando, assim, uma interpretação formal mais precisa.

Devido à nossa trajetória na área de Ciência da Informação com estudos específicos na área de Organização do Conhecimento e elaboração de linguagens documentárias, acreditamos que o presente trabalho venha oferecer uma contribuição teórico-metodológica consistente, no que diz respeito ao estabelecimento de padrões de relações entre conceitos.

Conforme o estudo foi sendo realizado, pôde-se perceber que as relações conceituais utilizadas pela Ciência da Informação para elaboração de seus vocabulários eram diferentes das relações que as ontologias necessitavam para possibilitar as inferências pela máquina. Essa percepção fez com que percorrêssemos outras literaturas, a fim de encontrarmos as relações necessárias para as Ontologias. Por isso, é importante ressaltar aqui, que além das bases teóricas que envolvem a Ciência da Informação e a Terminologia, essa pesquisa foi buscar em áreas como Ciência da Computação e Bioinformática o que já vinha sendo utilizado no que tange à utilização de relações. Sendo assim, este estudo vem trazer como contribuição uma sistematização do que essas quatro áreas vêm apresentando em termos de relações para elaboração de instrumentos de padronização terminológica.

O trabalho, daqui por diante, encontra-se organizado nos seguintes capítulos:

O segundo capítulo apresenta o objetivo geral e os objetivos específicos.

O terceiro capítulo apresenta uma revisão do que a literatura vem apresentando no que diz respeito à definição, estrutura, tipologia, aplicação e metodologia para construção de Ontologias, com o intuito de delimitar o objeto do estudo.

O quarto capítulo descreve a Gene Ontology, que é uma Ontologia de Domínio, utilizada para o estudo empírico.

O capítulo cinco apresenta uma revisão do que a literatura nas áreas de Ciência da Informação, Terminologia e Ciência da Computação vem apresentando no que diz respeito às relações conceituais, utilizadas em instrumentos de padronização terminológica.

O sexto capítulo apresenta uma revisão do que a literatura, na área de Bioinformática, vem desenvolvendo no que diz respeito também à utilização de relações conceituais no âmbito das ontologias.

O sétimo capítulo apresenta uma sistematização das relações na Ciência da Informação, na Terminologia, na Ciência da Computação e na Bioinformática.

No oitavo capítulo faz-se uma análise das definições em um corpus da Gene Ontology, aplicando-se o modelo de relações proposto.

No capítulo final, descrevem-se as conclusões obtidas através da pesquisa, as dificuldades encontradas e as perspectivas de estudos futuros.

2 OBJETIVOS

2.1 OBJETIVO GERAL

- Estudar as relações entre conceitos, identificar sua tipologia e verificar sua aplicação em Ontologias de domínio genômico.

2.2 OBJETIVOS ESPECÍFICOS

- Identificar as relações entre conceitos existentes na literatura da Ciência da Informação, da Terminologia e da Ciência da Computação.
- Estabelecer uma sistemática de relações entre conceitos.
- Identificar relações que podem ser aplicadas no âmbito de um corpus selecionado.
- Identificar, se for o caso, a necessidade de estabelecimento, nesse domínio, de outras relações ainda não identificadas na literatura.

3 ONTOLOGIA

As ontologias surgem na área de Inteligência Artificial, na década de 90, em uma fase chamada Moderna². Esta área tem por objetivo embutir conhecimento na máquina, através da intervenção humana. Em outras palavras, o que a Inteligência Artificial pretende é capacitar o computador para o entendimento de atividades que somente um ser humano seria capaz de efetuar.

Para os sistemas de Inteligência Artificial, o que existe é o que pode ser representado. Quando o conhecimento de um domínio é representado em uma linguagem declarativa, o conjunto de objetos que podem ser representados é chamado universo do discurso. Foi nesse sentido, que surgiram as ontologias, com o intuito de descrever programas, através da definição de um conjunto de termos que pudessem representar domínios e tarefas a serem executadas por estes programas.

Atualmente, o estudo de ontologias está voltado para a Web, onde se verifica um problema ainda muito comum: a recuperação imprecisa de informações. A solução dessa problemática tem atraído o interesse de profissionais das áreas de Computação e também de Informação, que acreditam que podem melhorar o nível de precisão das informações recuperadas, proporcionando mais semântica ao conteúdo das páginas Web.

² A Inteligência Artificial é dividida em três fases: Fase Clássica (56-70), que tinha por objetivo simular a inteligência humana através de métodos solucionadores gerais de problemas e lógica; Fase Romântica (70-80), cujo objetivo era simular a inteligência humana em situações pré-determinadas, através de formalismos de representação de conhecimentos adaptados ao tipo de problema, mecanismos de ligação procedural, visando maior eficiência computacional; Fase Moderna (80/90), cujo objetivo era simular o comportamento de um especialista humano ao resolver problemas em um domínio específico, através de um método de sistemas de regras, representação da incerteza, conexãoismo, segundo Classificação do Massachusetts Institute of Technology, apresentada em: <<http://www.das.ufsc.br/gia/history/>>.

Foi a visão da Web Semântica que re-introduziu os estudos de Ontologias na Comunidade Científica e, devido a isto, é importante mencionar como ocorreu essa retomada do estudo sobre ontologias por pesquisadores dessa área.

Na Web, assim como em bases de dados, para que haja um equilíbrio entre revocação e precisão, faz-se necessário, entre outros procedimentos, o controle do vocabulário, através da utilização de uma linguagem padronizada na realização do processo de atribuição de assunto do documento e também no momento da busca pelo usuário, para que desta forma se realize uma comunicação sem ruídos entre o sistema e o usuário.

Tim Berners-Lee, percebendo essa necessidade através de pesquisas desenvolvidas nos consórcios W3C - World Wide Web Consortium – propõe a Web Semântica.

A Web Semântica tem como visão a idéia de ter dados na Web definidos e ligados de uma maneira tal, que possam ser usados por máquinas, não só com o objetivo de apresentação, mas por automação, integração e reutilização de dados entre aplicativos (W3C, 2001).

Para que isto ocorra, é necessário um formalismo, que possa descrever propriedades e relacionamentos sobre itens, através de regras de inferências. Esta lacuna vem sendo preenchida através do uso de metadados e ontologias.

Enquanto os metadados (IANNELLA; WAUGH, 1997) descrevem as propriedades dos dados a serem representados, as ontologias ficam responsáveis pela padronização de significado, provendo a compatibilização de conceitos e a minimização dos problemas relacionados à comunicação, agora não apenas entre usuário e sistema, mas também entre sistema e sistema.

No contexto da Web Semântica, uma ontologia tem sido definida como um “documento ou um arquivo que define formalmente as relações entre os termos, sendo formada por uma taxonomia e regras de inferências” (BERNERS-LEE; HENDLER; LASSILA, 2001).

Ding (2001, p. 1) aborda as possibilidades de uso das ontologias na Web Semântica, a saber:

- Servir como dados sobre os dados para representação explícita da semântica desses dados em uma forma processável pela máquina.
 - Operacionalizar a semântica, oferecendo vários serviços inteligentes baseados em ontologias.
-

- Ajudar pessoas e computadores a acessar a informação que precisam.
- Ajudar computadores a se comunicarem efetivamente.

O uso de Ontologias na Web Semântica é considerado importante, pois permite que agentes de software entendam a semântica contida nas definições dos vocabulários de domínios específicos, diminuindo ambigüidades e propiciando o intercâmbio de informações, através das consultas às ontologias. Moreira, Alvarenga e Oliveira (2004, p. 137) definem agentes como “entidades de *software* que empregam técnicas de inteligência artificial com o objetivo de auxiliar o usuário na realização de uma determinada tarefa, agindo de forma autônoma e utilizando a metáfora de um assistente pessoal”. Além disso, Bases de Conhecimento também podem ser criadas, especializando e instanciando determinada ontologia, através de uma aplicação específica (MOURA, [2002]), isto é, definindo os elementos mais específicos.

As ontologias são importantes no âmbito da Web Semântica, por relacionarem automaticamente uma página com outras, através do emprego de regras de inferência, levando o usuário a novas informações (FERNEDA, 2003, p. 118); e por codificarem as páginas Web, permitindo que a informação seja interpretada pelos computadores, sem a necessidade de intervenção humana (TELLO, 2002).

O uso de Ontologias no âmbito da Web Semântica favorece ainda, “o compartilhamento da mesma estrutura de informações entre pessoas e softwares, pois possibilita a descrição formal das relações existentes entre os objetos em um formato que as máquinas possam identificar” e “a associação de uma ontologia a uma página Web definindo o significado de cada uma das informações existentes possibilita a integração e reutilização de ontologias entre diversos domínios” (RAMALHO, 2006, p. 59).

Na literatura levantada nas áreas de Ciência da Informação e da Ciência da Computação foi possível observar quatro tipos de abordagens de estudo sobre o assunto, a saber: 1) origem do termo, definição e composição das Ontologias; 2) possibilidades de aplicações de Ontologia; 3) tipos de Ontologias e 4) metodologias para construção de ontologias.

3.1 ORIGEM DO TERMO, DEFINIÇÃO E COMPOSIÇÃO DAS ONTOLOGIAS

O termo *Ontologia* vem do grego “*ón, óntos*”, que quer dizer “*ser*”, adicionado de “*logos*”, que é o *estudo de algo*. Sendo assim, a palavra *Ontologia* significa originalmente o estudo do ser. Atualmente, possui abordagem distinta na Filosofia e na Ciência da Computação.

O termo *Ontologia* é tomado como empréstimo da área de Filosofia pela área de Inteligência Artificial para designar um novo instrumento de tratamento e recuperação de informações.

Devido à variedade de literatura existente sobre o assunto, optou-se por fazer uma seleção dos autores levantados. O critério estabelecido para essa seleção foi o critério dos mais citados.

Corazzon (2000) faz uma longa revisão do que a literatura das duas áreas entendem por “Ontologia”. Sem nos determos nos aspectos filosóficos que, em nossa pesquisa, têm caráter secundário, registramos os filósofos citados por ele: Bolzano, Brentano, Meinung, Husserl, Cocchiarella, Poli. Na área da Ciência da Computação ele se limita a definições de Gruber e Guarino, que mais adiante são retomadas.

Algumas definições apresentadas na literatura de Ciência da Computação e principalmente na área de Inteligência Artificial, relacionam as *ontologias* a estruturas de conceitos.

Chandrasekaran, Josephson e Benjamin (1999, p. 20), por exemplo, afirmam que na Inteligência Artificial (AI), o termo *ontologia* está relacionado “com um vocabulário de representação, geralmente especializado em algum domínio ou assunto”, qualificado por conceituações de tipos de objetos e suas relações no mundo, ou em outras palavras com um “corpo de conhecimento que descreve algum domínio, usando um vocabulário de representação”.

Para Swartout e Tate (1999, p. 18) na Inteligência Artificial o termo ontologia é usado para se “referir a um conjunto de conceitos ou termos que podem ser usados para descrever alguma área do conhecimento ou construir uma representação dela”. Weinstein (1998, p. 256) define ontologias como uma “rede de definições de um vocabulário que expressa um consenso da comunidade sobre o domínio do conhecimento”, e Sowa (2000, p. 493), afirma que ontologia “é o estudo das categorias de coisas que existem ou podem existir no mesmo domínio”.

Da forma como essas definições foram formuladas, parece que uma ontologia é um vocabulário com uma lista de termos e definições, mas na verdade, as

ontologias são mais que isso, elas precisam de um algo mais que as torne processáveis por máquina.

Gruber (1993, p. 200) define ontologia como sendo “uma especificação formal e explícita de uma conceituação compartilhada”. Para compreender essa definição, outros autores recorrem ao texto de Ding e Foo (2002a, p. 123), onde estes procuram explicar o significado de cada um desses termos:

- **Conceituação** – “modelo abstrato de um fenômeno no mundo”;
- **Explícita** – “os tipos de conceitos usados e suas restrições devem estar explicitamente definidos”;
- **Formal** – “a ontologia deve ser processada por máquina”;
- **Compartilhada** – “a ontologia deve capturar o conhecimento aceito por consenso pelas comunidades que delas fazem uso”.

Gruber afirma a existência de uma parte formal nas ontologias e pelas definições dos elementos apresentadas por Ding e Foo, pode-se compreender que ele se refere a um tipo de linguagem que tornará a ontologia processável por máquina, além de também mencionar a existência de uma conceituação que precisa ser explicitada através de definições e compartilhada por sua comunidade de usuários.

Para Guarino, o uso mais freqüente do termo “ontologia” na área de Inteligência Artificial se refere a um “artefato de engenharia, constituído de um vocabulário específico usado para descrever uma certa realidade, mais o conjunto de pressupostos explícitos relacionados a um significado pretendido do vocabulário”. Para ele, num exemplo mais simples, “uma ontologia descreve uma hierarquia de conceitos relacionados pela reunião de relacionamentos” (GUARINO, 1998; GUARINO; GIARETTA; CARRARA, 1993).

Guarino (1998) afirma também que uma “Ontologia é uma teoria lógica que leva em conta o significado pretendido de um vocabulário formal, isto é, seu compromisso ontológico com uma conceituação particular do mundo”. Segundo ele, os modelos pretendidos de uma linguagem lógica que usa tal vocabulário são restringidos por seu compromisso ontológico: “Uma ontologia reflete indiretamente este compromisso (e a conceituação subjacente) aproximando estes modelos pretendidos”.

Smith (2002, p. 22) define “ontologia” na Ciência da Computação em relação a sistemas de informação como “artefato de software (ou linguagem formal), modelado com um conjunto específico de usos e ambientes computacionais em mente”.

Desta forma, o algo mais que as Ontologias precisam ter para assim serem consideradas é a linguagem formal, que permite que a máquina faça inferências automáticas sobre determinada questão.

O primeiro autor na área de Ciência da Informação a se ocupar do objeto Ontologia foi Vickery (1997), apoiando o conceito de ontologia nas definições dos autores da Ciência da Computação já mencionados neste estudo.

Em geral, pode-se verificar que na literatura da Ciência da Informação, os estudos sobre a definição do conceito de Ontologia estão voltados para uma tentativa de distinção entre os conceitos de Ontologia, Tesouro e Classificação, tentando assim, uma definição do objeto.

Pode-se citar como tentativas de distinção entre os conceitos de Ontologias e Tesouros nessa área, os textos de Moreira, Alvarenga e Oliveira (2004) e de Tristão, Fachin e Alarcon (2004).

Moreira, Alvarenga e Oliveira apresentam um artigo exaustivo de levantamento das definições em literaturas da Ciência da Informação e da Ciência da Computação e concluem, entre outras coisas, que tesouros e ontologias podem ser considerados instrumentos distintos, já que os tesouros servem de “instrumento de registro terminológico e para serem usados por pessoas, e não para registro do conhecimento para inferências computacionais” (MOREIRA; ALVARENGA; OLIVEIRA, 2004, p. 28). Para estes autores ainda, a diferença entre os instrumentos “pode ocorrer em termos de linguagem, de nível de formalização e de propósitos” (p. 29).

Tristão, Fachin, Alarcon (2004) apresentam tesouro e ontologias como instrumentos convergentes no que diz respeito às suas finalidades, ou seja, organização do conhecimento e recuperação da informação.

Toda definição de Ontologia utilizada na área de Ciência da Informação é buscada em uma revisão de literatura da área de Ciência da Computação. Apenas alguns autores, mais recentemente, tentam uma nova definição do termo. Para CAMPOS (2001a, p. 109) ontologia é “um conjunto de conceitos padronizados,

termos e definições aceitos por uma comunidade particular”. Para Ramalho (2006, p. 59) ontologias são:

instrumentos de representação do conhecimento definidos em uma linguagem formal e processável por máquina que possibilitam a descrição dos aspectos semânticos dos conteúdos informacionais, explicitando seus relacionamentos de modo detalhado a partir de restrições lógicas que possam ser processadas de forma automatizada, possibilitando inclusive relacionamentos baseados na Lógica de Segunda Ordem.

Com esta última definição, mais uma vez, pode-se perceber a necessidade das ontologias serem processáveis por máquina.

Sobre a definição do objeto “Ontologia”, este estudo destaca os seguintes pontos em relação às definições apresentadas:

- Ontologias possuem conceituações que devem ser compartilhadas.
- Ontologias devem ser descritas através de axiomas lógicos.
- Ontologias devem ser processáveis por máquina.

Falar que as ontologias devem possuir conceituações compartilhadas é dizer que elas devem possuir um vocabulário estruturado com termos, definições e relacionamentos, que devem expressar um acordo comum entre seus usuários.

Falar que elas devem ser descritas através de axiomas lógicos é dizer que suas definições devem possuir uma forma padrão, baseada em axiomas, que possam determinar a verdade das sentenças.

Falar que elas devem ser processáveis por máquinas é dizer que, além de um vocabulário, a ontologia deve possuir também uma linguagem formal que propicie a interpretação desse vocabulário pela máquina, ou seja, uma linguagem formal. Sendo assim, as ontologias têm como componentes, que fazem parte de sua estrutura, os seguintes elementos (TELLO, 2004):

Conceitos - que são idéias básicas sobre o que se pretende formalizar. Os conceitos podem estar organizados em classes de objetos, métodos, planos, estratégias, processos etc.

Classes e Subclasses - que podem formar uma taxonomia.

Relações - que devem representar os tipos de interação entre as classes de um domínio. Essas relações são formalmente definidas como qualquer subconjunto dos produtos de um conjunto e são sempre binárias, como por exemplo: `subclasse_of`, `connected_to` etc.

Funções – são casos especiais de relações no qual os elementos dos relacionamentos são únicos para os elementos anteriores, por exemplo a relação mother_of. Em outras palavras são relações que não permitem inversa.

Axiomas - são teoremas que se declaram sobre as relações que devem cumprir todos os elementos da ontologia.

Instâncias - são utilizadas para representar objetos determinados de um conceito.

Todos estes elementos devem possuir uma representação formal para que a ontologia seja processável pela máquina e possa atingir o seu objetivo, que é promover comunicação entre pessoas, organizações e/ou sistemas de software (USCHOLD; GRUNNINGER, 1996, p. 93).

3.2 POSSIBILIDADES DE APLICAÇÃO DE ONTOLOGIAS

Quanto às possibilidades de aplicação de ontologias, muitos são os autores que abordam a questão e também diversas as formas de uso deste instrumento. Nesta seção, procura-se reunir as possibilidades de aplicação de forma sistematizada a partir da literatura, segundo as seguintes categorias:

- Comunicação
- Especificação de Sistemas,
- Reutilização,
- Engenharia de sistemas
- Confiabilidade
- Processamento de textos em Linguagem Natural
- Classificação para ação
- Usabilidade
- Interoperabilidade
- Representação de Informação
- Recuperação de Informação

Na categoria “Comunicação”, “as ontologias são usadas com objetivo de reduzir a confusão terminológica e conceitual através do provimento de uma unificação do quadro terminológico de uma organização” (USCHOLD; GRUNNINGER, 1996). Desta forma, as ontologias podem compartilhar entendimento e comunicação

entre pessoas com necessidades e pontos de vistas distintos, a exemplo de modelos normativos. Para Vickery (1997), as ontologias permitem que diferentes léxicos possam comunicar o conhecimento através de uma linguagem única padrão.

Na categoria “Especificação do Sistema”, as ontologias apresentam regras que podem variar de acordo com o nível de formalidade e automação dentro da metodologia de modelagem do sistema, facilitando assim, o processo de identificação de requisitos do sistema e a compreensão de relacionamentos entre seus componentes (USCHOLD; GRUNINGER, 1996; VICKERY, 1997; JASPER; USCHOLD, 1999; MOREIRA, 2002).

Na categoria “Reutilização”, as ontologias podem auxiliar na importação e exportação de módulos em diferentes domínios. As ontologias devem permitir uma facilidade para re-uso de classes de objetos, modelagem de problemas e domínios. Para isso, as ontologias devem ser construídas de forma padronizada. O re-uso de ontologias propicia que uma linguagem possa ser traduzida em outras linguagens e utilizada em outras aplicações (USCHOLD; GRUNINGER, 1996; SWARTOUT et al., 1997; JASPER; USCHOLD, 1999).

Na categoria “Engenharia de Sistemas”, as ontologias são usadas para apoiar a projeção e o desenvolvimento de sistemas de software (USCHOLD; GRUNINGER, 1996).

Na categoria “Confiabilidade”, as ontologias são utilizadas para prover consistência ao software. Neste sentido, podem ser utilizados dois tipos de ontologias: ontologias informais e ontologias formais³. Enquanto as ontologias informais servem à confiabilidade dos software, agindo como um manual de verificação do modelo utilizado para a especificação; as ontologias formais possibilitam o uso de software de verificação de consistência (semi)automático com respeito à especificação declarativa, além de poderem ser usadas para facilitar a integração entre sistemas, já que tornam explícitas as várias suposições feitas pelos diferentes componentes de um sistema. Em outras palavras, existem duas formas da ontologia auxiliar na Confiabilidade do sistema: uma é ter um vocabulário em linguagem natural (ontologia informal) acoplado no sistema, para ser consultado pelos programadores e outra é utilizar a ontologia formal, que estará acoplada no sistema, em linguagem formal, possibilitando que o próprio software faça as

verificações do que está certo ou errado, em termos de especificações (USCHOLD; GRUNINGER, 1996).

Na categoria “Processamento de textos em linguagem natural”, uma ontologia provê que a extração de conhecimento de textos em linguagem natural seja realizada de forma eficaz, tendo em vista que a ontologia padroniza os significados dos termos utilizados para extração (VICKERY, 1997; ALMEIDA; BAX, 2003).

Na categoria “Classificação para ação”, as ontologias permitem que as tarefas a serem executadas sejam classificadas, possibilitando estabelecimento de prioridades para cada tarefa, de acordo com o tipo de ação a ser realizada, além de permitir a padronização das tarefas e o compartilhamento de uma linguagem de execução comum a softwares distintos (SOERGEL, 1999).

Na Categoria “Usabilidade”, Moreira (2002) afirma que as ontologias podem “prover interfaces corporativas” e “auxiliar na navegação em bases de informação”. Isso quer dizer que as ontologias podem padronizar o significado dos termos utilizados em interfaces corporativas e também auxiliar através de mecanismos de classificação a arquitetura da informação nas bases de informação, propiciando desta forma, uma interação mais amigável para o usuário.

O uso de Ontologias para promoção de Interoperabilidade entre sistemas é mencionado pela maioria dos autores e pode ser considerada uma outra categoria, apesar de algumas vezes aparecer com uma nomenclatura diferente. Neste sentido, os autores afirmam que o uso das Ontologias se faz necessário em situações em que se têm “diferentes usuários que necessitem trocar dados e utilizar diferentes softwares” e citam como exemplo de uso de ontologias, objetivando a interoperabilidade, as ontologias como inter-língua - que visam, entre outras coisas, a integração de modelos, de repositórios de dados e também o uso para suporte de traduções entre diferentes linguagens e representações (USCHOLD; GRUNINGER, 1996; VICKERY, 1997). Por exemplo, na área de Bioinformática, a integração de Bases de dados heterogêneas é fundamental, tendo em vista que promovem a “interoperabilidade entre diferentes ferramentas de análises e algoritmos mais sensíveis para detecção de homologias mais distantes”. (MENDES, 2006, p.61)

³ A diferença entre Ontologias Informais e Ontologias Formais está no fato da primeira ser expressa apenas em linguagem natural, enquanto a segunda é definida também através de linguagem de inferência.

Na categoria “Representação de Informação”, as ontologias são usadas a fim de prover uma base para representação do significado do texto, melhorando assim a comunicação e a recuperação de informações (SOWA, 2000; VICKERY, 1997).

3.3 TIPOS DE ONTOLOGIAS E SUAS APLICAÇÕES

Inúmeras são as tipologias registradas na literatura, algumas delas conflitantes. Em seu artigo de revisão, Almeida e Bax (2003) classificam os diversos tipos, segundo o critério dos cinco autores estudados, a saber:

- quanto à função
- quanto ao grau de formalismo
- quanto à aplicação
- quanto à estrutura
- quanto ao conteúdo

Segundo os critérios de tais autores, alguns tipos se repetem, como é de se esperar. Devido a isto, este estudo foi levado a propor uma nova classificação dos tipos de Ontologias. A intenção desta classificação não foi agrupar toda a tipologia conforme mencionada na literatura, mas expor de forma mais precisa a tipologia de ontologias que seria útil para o desenvolvimento desta pesquisa.

Uma revisão mais ampla da literatura permitiu produzir uma classificação que levou em conta o princípio da exclusividade, ou seja, cada tipo de ontologia está classificado em uma e única classe.

Para a compreensão da importância das relações conceituais para a elaboração de ontologias, foi importante apresentar a classificação das ontologias de acordo com dois critérios: por sua natureza e pelo grau de formalismo.

A classificação pela natureza das ontologias reuniu as Ontologias de Domínio e as Ontologias de Tarefa. Já a classificação pelo grau de formalismo reuniu as ontologias informais, lingüísticas ou terminológicas, (ontologias semi-informais), ontologias formais, ontologias rigorosamente formais. Ver Figura 1:

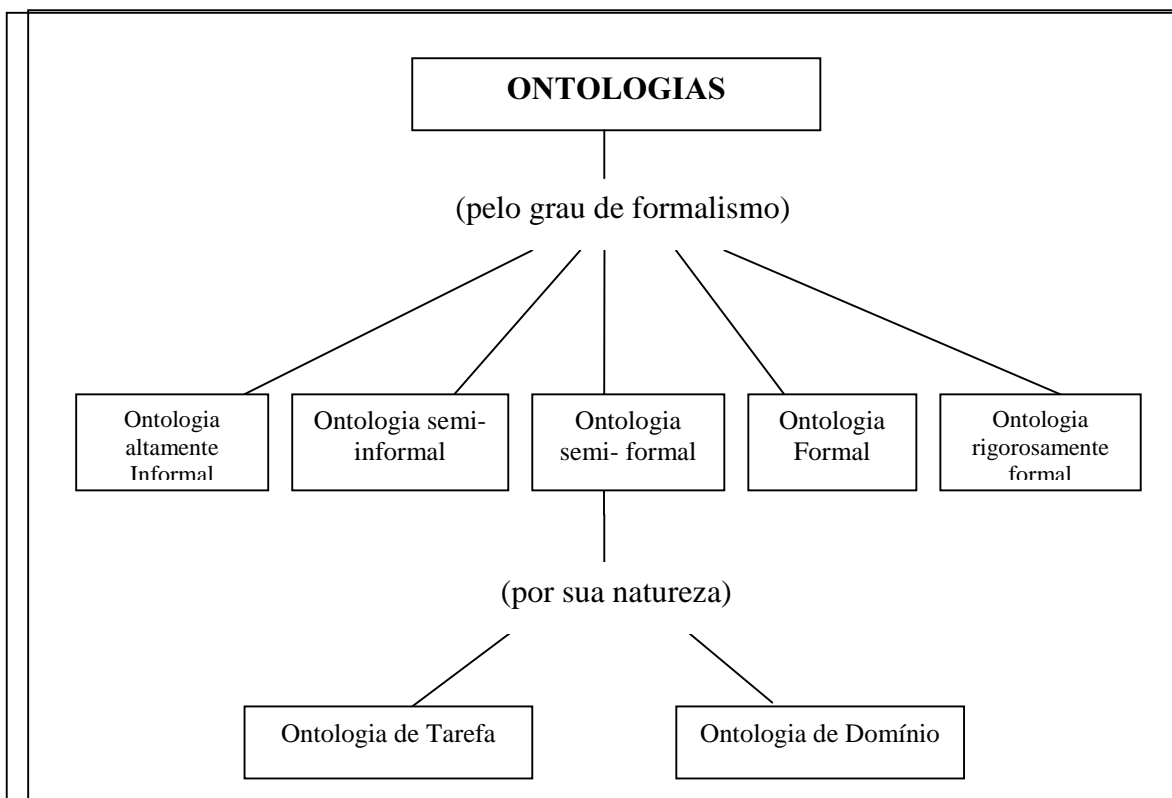


Figura 1 – Tipos de Ontologias

➤ Ontologias: quanto ao grau de formalismo

As Ontologias podem ser classificadas quanto ao grau de formalismo em: altamente informais, semi-informal, semi-formal, formal e rigorosamente formal.

Esta pesquisa utiliza as definições de ontologias quanto ao grau de formalismo dadas por Uschold e Gruninger (1996), com exceção apenas para Ontologia Formal, que será assumida neste trabalho como um tipo de ontologia que está entre a ontologia semi-formal e a ontologia rigorosamente formal, isto é, ela utiliza vocabulário controlado em linguagem natural e a partir dele constrói o seu vocabulário em linguagem artificial, com base na lógica de primeira ordem.

As ontologias Altamente Informais, segundo Uschold e Gruninger (1996), são aquelas expressas livremente em linguagem natural, sem nenhum controle.

Van Heijst, Schreiber, Wielinga (1997) chamam as Ontologias Informais de Terminológicas e as definem como aquelas que especificam os termos que serão usados para representar o conhecimento em um domínio (por exemplo, os léxicos).

As Ontologias Semi-informais e Semi-formais só são definidas por Uschold e Gruninger (1996). Para eles, as primeiras são ontologias expressas em linguagem

natural de forma restrita e estruturada e as segundas são ontologias expressas em uma linguagem artificial definida formalmente.

As Ontologias Formais são definidas por Guarino (1998) como ontologias que utilizam a teoria da lógica de primeira ordem, nas quais as palavras do vocabulário aparecem como predicados unários ou binários, chamados respectivamente de conceitos e relações.

As Ontologias Rigorosamente Formais, segundo Uschold e Gruninger (1996), são aquelas expressas por termos que são definidos com semântica formal, teoremas e provas.

➤ Ontologias: por sua natureza.

Quanto à natureza das Ontologias, elas podem ser classificadas em: Ontologias de Tarefas e Ontologias de Domínio.

As Ontologias de Tarefa são aquelas que fornecem um vocabulário sistematizado de termos, especificando tarefas genéricas que podem ou não estar no mesmo domínio, como por exemplo, diagnóstico, vendas etc (MIZOGUCHI; VAN WELKENHUYSEN; IKEDA, 1995; GUARINO, 1998; HAAV; LUBI, 2001). Essas ontologias, portanto, descrevem atividades através de um vocabulário formado por ações com a finalidade de possibilitar a execução de tarefas pela máquina.

As Ontologias de Domínio são ontologias reutilizáveis em um domínio genérico e que fornecem vocabulários sobre conceitos, seus relacionamentos, sobre atividades e regras que os governam, ex. medicina, automóveis etc. (MIZOGUCHI; VAN WELKENHUYSEN; IKEDA, 1995; GUARINO, 1998; VAN HEIJST; SCHREIBER; WIELINGA, 2002). Esse tipo de Ontologia é próprio para representação e recuperação das informações de um dado domínio de conhecimento.

Guarino (1998, p. 4) define as Ontologias de Domínio em oposição às Ontologias de Tarefa. Segundo ele, as Ontologias de Domínio, diferentemente, das Ontologias de Tarefa, que expressam métodos e ações, são construídas através da especificação de conceitos de um dado domínio do conhecimento.

Na Ciência da Computação o estudo do domínio tem sido de grande interesse para o desenvolvimento de software. Para esta área, o termo “Domínio” pode ser

definido de vários modos, por exemplo, Berard (1992 apud HJORLAND, 2002) faz duas caracterizações:

(1) Coleção de aplicativos correntes e futuros (software) que compartilham um conjunto de características comuns.

(2) Conjunto bem definido de características que descrevem de modo acurado, específico e completo uma família de problemas, pelos quais as soluções dos aplicativos de computador são e serão buscadas.

Neste estudo abordaremos “domínio” como o conteúdo de uma área do conhecimento específica ou o seu campo de assunto. Nele também faremos uma análise de uma ontologia de domínio específica, isto é: a Gene Ontology, detalhada no capítulo 4.

3.4 METODOLOGIAS PARA CONSTRUÇÃO DE ONTOLOGIAS

A literatura assinala a ausência de uma metodologia unificada e respaldada por bases teóricas como sendo um dos maiores problemas encontrados no âmbito do desenvolvimento de ontologias. Alguns artigos são bastante detalhados, mas apenas os aspectos relacionados à pesquisa são enfatizados.

Bouaud e outros (1994, 1995) apresentam um sistema chamado MENELAS que possui quatro princípios para modelagem de ontologias a partir de um sistema de entendimento de linguagem natural. Estes princípios estão relacionados à modelagem de ontologias a partir da elaboração de taxonomias. e se referem a similaridade e especificidade.

No que diz respeito à similaridade, uma subclasse deve ser do mesmo tipo que seu pai. No que diz respeito à especificidade, a subclasse deve ter alguma diferença que a distingue do pai. Uma diferença como definição de um pai é uma forma necessária e uma condição suficiente para a definição da subclasse. A adoção de tais princípios deve fornecer as condições para o estabelecimento de hierarquias.

Gruninger e Fox (1995) propõem uma metodologia chamada TOVE (Toronto Virtual Enterprise). A questão das relações é mencionada na segunda e na terceira etapa desta metodologia. A segunda etapa, os autores chamam de “Informal Competency Question”, onde se identifica a motivação do cenário através da seleção das questões que a ontologia deverá responder. Para isso, sugerem uma

série de questões que devem ser respondidas sobre o contexto no qual a ontologia deve ser construída. É nesta etapa que sugerem que seja feita a definição da terminologia da ontologia, ou seja, seus objetos, atributos e relações. Esta primeira definição ainda é feita em linguagem natural. A terceira etapa, chamada de “Especificação da terminologia em lógica de primeira ordem”, deve conter a especificação de toda a terminologia definida na etapa anterior, em lógica de primeira ordem.

Uschold e King (1995) criam uma ontologia chamada ENTERPRISE e abordam a questão das relações na etapa de Captura do Conhecimento. Nesta etapa é realizada a identificação de conceitos chaves e dos relacionamentos de um domínio; e também a produção precisa de definições textuais sem ambigüidades para cada conceito e seus relacionamentos.

Fazendo uma revisão da literatura sobre o tema, Uschold (1996) verifica que, em geral, as metodologias se limitam a estudos de caso do desenvolvimento de uma única ontologia, ou a um projeto em particular. Propõe então uma ‘metodologia unificada’, ou seja, derivada das metodologias TOVE e ENTERPRISE, que procuram identificar as etapas e técnicas de aplicabilidade geral; identificar as circunstâncias em que se aplicam etapas e técnicas não-gerais, tentando colocar tudo num quadro coerente.

Nesta metodologia unificada, as etapas não estão distintas claramente, mas ele fala de algumas ações realizadas no processo anterior à construção da ontologia, que podem ser úteis na identificação do vocabulário e na produção de definições, a saber:

- Identificar o propósito
- Identificar o nível de formalidade necessário
- Identificação do escopo – nível de assunto – A produção dessa fase é um conjunto de conceitos e termos cobrindo toda gama de informação que a ontologia deve caracterizar para satisfazer as exigências já identificadas
- Construção da ontologia
- Avaliação/Ciclo da revisão
- Acompanhamento

Nas etapas da construção, a que merece destaque é a segunda, sobre a produção de Guias para criação das definições, que devem apresentar:

- **Clareza** – as definições devem ser claras e sem ambigüidades, se expressas em linguagem natural, ou formalmente codificadas, usando exemplos sempre que possível a fim de ilustrar o que se pretende.
- **Consistência e Coerência** – Uma ontologia deve ser internamente consistente, a circularidade deve ser evitada, especialmente se deseja codificação formal. Deve-se evitar introduzir neologismos, e consultar dicionários, tesouros e glossários técnicos.
- **Extensibilidade e Reusabilidade** – uma ontologia deve ser modelada de tal forma que seja possível o máximo de reutilização e extensão e extensibilidade seja possível. Deve-se evitar introdução de diversos termos que signifiquem aproximadamente a mesma coisa.
- **Go Middle-out** - Após escolha e definição dos termos, escolhe-se a abordagem de estruturação dos termos, ou seja, se top-down, bottom-up ou middle-out. O autor sugere o uso da abordagem middle-out, que significa estruturar os conceitos iniciais, introduzindo termos mais gerais e mais específicos, conforme a necessidade.
- **Lidando com ambigüidades** – Para conseguir concordância quando os termos são usados de forma ambígua, deve-se concentrar primeiro nas idéias subjacentes, ignorando os termos. Definir cada idéia relacionada, criando etiquetas sem significação para cada uma delas, depois decidir pela idéia mais importante e, finalmente, selecionar os termos apropriados.

O KBSI IDEF5 (SLATERY, 1997) é uma metodologia desenhada para criação, modificação e manutenção de ontologias e é dividida em cinco passos - Definição do escopo, Coleta de dados, Análise dos dados, Desenvolvimento e Refinamento - dos quais destacamos o quarto, em que uma ontologia preliminar é desenvolvida, com conteúdos de modelos de conceitos, isto é, descrições de tipos, relações e propriedades.

Neste sentido, Jones, Bench-Capon e Visser (1998) afirmam que “a construção de ontologias ainda é mais artesanal que científica” e que tem havido um “crescimento do número de metodologias específicas para desenvolvimento e manutenção de ontologias”, o que também se pode constatar através das diversas metodologias analisadas por eles.

A Infosleuth de Hwang (1999) trata de uma descrição de construção automática de ontologia a partir de uma base de dados textual e não está no âmbito deste estudo, porém destacamos algumas etapas, na qual essa construção automática utiliza uma classificação automática. Essa metodologia é dividida em cinco etapas : fornecimento de palavras por especialistas, extração automática de palavras dos documentos, classificação dos conceitos, verificação da classificação, expansão da ontologia , das quais destacamos duas: a etapa da classificação e da verificação.

A etapa da classificação⁴ é onde os conceitos da ontologia são classificados automaticamente por um software. Durante este processo, a ontologia também pode coletar palavras candidatas que estavam em volta do processamento.

A etapa da verificação é aquela em que um especialista verifica a classificação realizada automaticamente na etapa anterior e faz as correções necessárias.

A ABC Model de Brickley, Hunter e Lagoze (1999) descreve alguns princípios para representação de uma ontologia em RDF, fala das relações, que são típicas da modelagem em RDF, porém não define uma metodologia para construção de uma ontologia no todo, limitam-se à parte de modelagem em RDF. As Etapas são as seguintes:

1ª Etapa - Separar os termos em categorias (temporalidade, abstração e tempo)

2ª Etapa - Separar os conceitos em classes (que são os conceitos genéricos ou subcategorias).

3ª Etapa - Relacionar essas classes através das propriedades, que são preestabelecidas.

Kietz, Maedache e Volz (2000) criam uma metodologia chamada OntoKnowledge, na qual sugerem uma etapa para a “aquisição dos conceitos de domínios específicos de recursos disponíveis [...] para a base da ontologia” e também outra etapa para a “seleção das relações não-taxionômicas, ampliando assim o número de relações”. As relações taxionômicas, estranhamente, não são mencionadas, mas é importante destacar a menção, por estes autores, da

⁴ Apesar de acharmos que essa etapa não seja possível, seria preciso avaliar o software para verificar a consistência da classificação realizada pelo mesmo.

necessidade de outros tipos de relações que devem ser selecionadas em textos para compor as ontologias.

Noy e McGuinnes (2001) apresentam antes da descrição da metodologia algumas regras, que segundo eles, podem ajudar na tomada de decisões em muitos casos. Entre estas regras está uma que prescreve que os “conceitos na ontologia devem ser limitados a objetos (físicos ou lógicos) e também aos relacionamentos no domínio de interesse”. Essa metodologia parece ser a mais coerente, pois trabalha com sistemas de conceitos.

Sugerem que seja feita a “definição das classes e das hierarquias de Classes”. Segundo estes autores, existem várias possibilidades de desenvolver hierarquias de conceitos, a saber:

- **Top-down** – na qual o processo parte da definição dos conceitos mais genéricos e a subsequente especialização dos conceitos.
- **Bottom-up** – nesta, o processo começa pela definição de conceitos mais específicos do nível hierárquico, com o agrupamento subsequente de grupos de classes em conceitos mais gerais.
- **Combinação dos dois processos** – na qual primeiramente se define os conceitos mais salientes e em seguida generaliza e especializa os mesmos de forma apropriada.

Sugerem também que seja realizada a definição das propriedades das classes (slots)⁵. Para eles, em cada propriedade da lista devem ser determinadas quais classes elas descrevem. Estas propriedades se tornam slots anexados a classes e em geral existem diversos tipos de propriedades de objetos que podem tornar slots nas ontologias:

- Propriedades Intrínsecas – propriedades como o sabor do vinho
- Propriedades Extrínsecas – propriedades tais como, o nome do vinho ou o lugar de onde ele vem etc.
- Partes – se o objeto é estruturado, as “partes” podem ser tanto físicas como abstratas.
- Relacionamentos com outros indivíduos – são relacionamentos entre membros individuais das classes e outros itens.

⁵ Slots são algumas vezes usados como sinônimos de propriedades/atributos, mas os slots descrevem relacionamentos.

Outra sugestão é a definição das facetas dos slots. Esses slots podem ter diferentes facetas, descrevendo o tipo de valor, o número de valores (cardinalidade), e outras características de valores que o slot pode ter.

Como pode ser observado, este autor se refere a propriedades, facetas e relacionamentos, o que faz com que, implicitamente, ele esteja se referindo a uma estrutura de conceitos.

Sure, Staab e Studer (2002) apresentam uma metodologia para criação e manutenção de uma ontologia baseada na aplicação da gestão do conhecimento nas empresas. Esta metodologia é dividida em cinco etapas – Viabilidade do Estudo, Partida, Refinamento, Avaliação e Aplicação e Evolução - e a questão das relações é tratada em duas delas.

Na etapa de partida, a ontologia começa efetivamente a ser criada. Nela é criado um documento de especificação do conhecimento, que descreve o que a ontologia deve apoiar, esboçando o plano do âmbito de aplicação da ontologia e listando, por exemplo, recursos de conhecimento valiosos para reunir a descrição semi-formal da ontologia. Este documento deve auxiliar o engenheiro de ontologia a decidir sobre inclusão e exclusão de conceitos, e relações na estrutura hierárquica da ontologia. O resultado desta fase é uma descrição semi-formal da ontologia. Esta é uma etapa de tomada de decisão.

A etapa de Refinamento é aquela durante a qual pode-se escolher entre a metodologia bottom-up ou top-down para refinamento dos conceitos. Nesta etapa também se formaliza a descrição inicial da ontologia semi-formal. Primeiramente, os engenheiros formam uma taxonomia a partir da descrição semi-formal da ontologia e adicionam outras relações além de "is_a", às quais formam uma estrutura taxonômica.

Almeida (2003, p. 175) propõe um roteiro para elaboração de Ontologias, sugerindo etapas a serem seguidas na sua construção. A questão das relações também são mencionadas, mas ele aborda com mais ênfase a definição das classes, da hierarquia e das propriedades.

Pinto e Martins (2004) apresentam uma metodologia para construção de ontologia, caracterizada pelo ciclo de vida de uma engenharia de software. Geralmente, os estágios aceitos para construção de ontologia são especificação, conceituação, formalização, implementação e manutenção. Em cada um desses estágios os autores apresentam atividades. Aqui será destacado o estágio das

atividades que abordam a questão da conceituação, da formalização e da implementação, onde tratam das definições e das relações.

No estágio da Conceituação deve-se descrever, num modelo conceitual, a ontologia a ser construída. Diferentes metodologias propõem o uso de diferentes modelos conceituais: formal, informal, semi-formal etc. O modelo conceitual de uma ontologia consiste em conceitos do domínio e relacionamentos entre esses conceitos e esses relacionamentos melhoram a força das conexões entre os conceitos.

No estágio da Formalização, a descrição conceitual é transformada em um modelo formal, isto é, a descrição do domínio encontrada no passo anterior é escrita em uma linguagem própria para máquina. Nesta etapa, conceitos são definidos através de axiomas que restringem as interpretações possíveis para o significado desses conceitos. Estes conceitos são geralmente organizados hierarquicamente, através de uma relação de estruturação, tais como *é_um* (classe e super-classe, instância-classe) ou *parte de*.

No estágio da Implementação, a ontologia é formalizada em uma linguagem de representação do conhecimento. Para isso, compromete-se com um modelo formal de linguagem de representação.

Além do estudo das relações propriamente ditas, a análise do que as metodologias vêm falando sobre a questão das definições e taxonomias é considerada importante, já que muitas das vezes, as relações conceituais não aparecem explicitamente nas terminologias, mas implicitamente dentro das próprias definições. Além disso, as taxonomias muitas das vezes são usadas como ponto de partida para o estabelecimento de relações, já que através das classificações são reveladas relações *Is_a* e/ou *part_of*, dependendo da metodologia em questão.

Na área de Ciência da Informação, como era previsto, não se encontram metodologias específicas para construção de Ontologias, já que estas não são seu objeto de estudo, porém possui metodologias, com bases teóricas consistentes, para construção de Linguagens Documentárias, que sugerem relações entre conceitos.

Autores como Campos (2000a), Moreira e Oliveira (2005) assinalam uma possível contribuição das bases teóricas que regem a construção de Linguagens Documentárias para a construção de ontologias. Esta pesquisa compartilha da opinião destes autores, tendo em vista que na área de Ciência da Informação e mais precisamente na área de Organização do Conhecimento, podem-se encontrar metodologias bem estabelecidas para a elaboração de classificações

(RANGANATHAN, 1967), de definições (DAHLBERG, 1978a) e de relações entre conceitos (WÜESTER, 1981; FELBER, 1984). Desta forma, a Ciência da Computação pode se beneficiar através do preenchimento dessas lacunas existentes no âmbito da construção de ontologias, com a base teórica da Ciência da Informação.

Esta pesquisa visa sistematizar as possibilidades de relações apresentadas na literatura, a fim de prover a área de elaboração de ontologias de um modelo para estabelecimento das mesmas. Autores como Wüster (1981), Felber (1984) e Dahlberg (1978) apresentam uma classificação das relações entre conceitos existentes, mas isso será tratado de maneira mais detalhada no capítulo 5.

4 A GENE ONTOLOGY: UMA ONTOLOGIA DE DOMÍNIO

Uma Ontologia de Domínio é um tipo de ontologia própria para representação de conceitos em uma determinada área de assunto. Outras definições encontradas na literatura apresentam aspectos adicionais para compreensão deste objeto.

Swartout e outros (1996, p. 2) afirmam que as Ontologias de Domínio devem “prover um conjunto de termos para a descrição de um domínio” e podem “ser pensadas como provedoras de uma taxonomia de objetos relevantes de um mesmo domínio”. Para este autor, as Ontologias de Domínio possuem milhares de conceitos e tendem a ser grandes e por isso requerem um bom conhecimento do domínio, além da participação de especialistas no seu desenvolvimento.

Guarino (1997, p. 11) considera as Ontologias de Domínio como relativamente estáticas, pois as partes genéricas mudam quando a estratégia de solução de problemas também muda.

Para Musen (1998, p. 540), Ontologias de Domínio são “descrições formais de classes de conceitos e dos relacionamentos entre esses conceitos que descrevem uma área de aplicação”.

Quanto ao uso, elas servem para “gerar ferramentas de aquisição de conhecimento em um domínio específico, podendo facilitar a modelagem de domínios” (MUSEN, 1998, p. 553). Sendo assim, as Ontologias de Domínios podem também ser um instrumento para entendimento/compreensão de domínios desconhecidos.

Um outro exemplo de uso de Ontologias de Domínio é a elaboração de um índice com significado dos documentos, onde através de suas definições, elas possam auxiliar na compreensão dos conceitos de um texto e no uso de conhecimento por vários aplicativos (WU; TSAI; HSU, 2003, p. 33). É possível

também fazer uso das Ontologias de Domínio para a execução de uma aplicação específica, como por exemplo, anotar documentos, encontrar itens etc. (VAN ELST; ABECKER, 2001, p. 6), como na área de Bioinformática.

Quanto à necessidade das Ontologias de Domínio, pode-se afirmar, de acordo com Navigli e Velardi (2004, p. 1), que a “importância das Ontologias de Domínio é amplamente reconhecida, particularmente, em relação ao advento da Web Semântica”. Para estes autores, o objetivo principal de uma Ontologia de Domínio é “reduzir (ou eliminar) a confusão conceitual e terminológica entre membros de uma comunidade virtual de usuários... que necessitam compartilhar documentos eletrônicos e informação de vários tipos”.

Apesar do amplo reconhecimento do aumento do número de domínios e, conseqüentemente, o aumento da necessidade de uso de Ontologias de Domínio, diversas barreiras ainda devem ser vencidas antes delas se tornarem ferramentas utilizadas, como por exemplo, torná-las processáveis, mas para isso, é preciso ter relações conceituais bem definidas.

O presente trabalho utiliza como campo empírico, para análise das relações, uma Ontologia específica do domínio Genômico: A Gene Ontology (GO)⁶.

A Gene Ontology faz parte de um grupo de Ontologias chamado de OBO⁷ (Open Biomedical Ontologies), que tem por objetivo promover uso compartilhado entre diversos domínios biológicos e médicos. Segundo Mendes (2005, p. 30), nesse grupo de ontologias “alguns vocabulários são mais genéricos, pois objetivam ser aplicáveis a quaisquer organismos, outros são mais específicos, pois objetivam representar grupos taxonômicos específicos, tais como moscas, fungos, leveduras ou peixes”.

Apesar de existirem diversas ontologias e diversos grupos inserindo dados na OBO, a inclusão de dados em qualquer uma das ontologias pertencentes ao grupo é feita de acordo com critérios que devem ser seguidos fielmente, ou os termos não são aceitos pelo consórcio. Como a Gene Ontology também faz parte do grupo, esses critérios também são utilizados para inserção de termos, a saber:

- A ontologia deve ser aberta;

⁶ Cf. <<http://www.geneontology.org/>>.

⁷ Cf. <<http://obo.sourceforge.net/>>.

- Deve usar a sintaxe OWL;
- O termo deve ter um identificador único (*unique identifier space*); e
- Deve incluir definição textual de seus termos.

Com esses critérios, o consórcio OBO visa a ter um vocabulário controlado para: escrever e manter as ontologias; fazer associações entre as ontologias; e desenvolver ferramentas que facilitem a criação, manutenção e uso de ontologias, promovendo assim, uma troca de dados mais precisa.

Como as ontologias da OBO seguem um padrão preestabelecido, poderia ter sido escolhida qualquer outra ontologia pertencente ao grupo, no entanto, a escolha da Gene Ontology ocorreu devido ao fato desta ontologia ser utilizada no grupo Dataware, coordenado pela Prof. Dr. Maria Luiza Machado Campos da Pós Graduação em Ciência da Computação da UFRJ, do qual temos participado⁸ em atividade de pesquisa interdisciplinar, desde a graduação.

Uma das vertentes de pesquisa do grupo Dataware é voltada para tratamento e recuperação de informação como fator estratégico para auxiliar a pesquisa básica no âmbito da Bioinformática. Especificamente, na área de aplicações científicas, vem desenvolvendo o projeto “Genoma e Transcriptoma comparativo, um consórcio de Bioinformática para o desenvolvimento de uma plataforma Web e bancos de dados integrados”, atualmente financiado pelo CNPq e coordenado pelo Dr. Alberto M. R. Dávila da FIOCRUZ. Este projeto tem como um dos principais objetivos prover um ambiente que possa oferecer informação semântica sobre recursos científicos, como dados e programas, na área de Bioinformática e possibilitar o uso destes recursos de forma conjunta, pela comunidade científica interessada. Para isso, é essencial o uso de ontologias enquanto instrumento de padronização terminológica, próprio para troca de informação.

A Ontologia utilizada pelo grupo é a Gene Ontology, que produz um vocabulário controlado para descrição de genes e atributos de produtos genéticos de alguns organismos.

⁸ A participação nesse grupo de pesquisa se deu por intermédio da Prof. Dra. Maria Luiza de Almeida Campos, da Universidade Federal Fluminense, nossa orientadora desde os estudos de graduação (Projetos PIBIC e TCC) e que, na época, também atuava como professora convidada do Programa de Pós-Graduação em Ciência da Computação da UFRJ.

Esta Ontologia é amplamente utilizada na área Biomédica, mas seu uso mais intenso é na área de Genoma. Com ela, os pesquisadores fazem anotações de informações importantes, identificadas no momento da análise da comparação de seqüências.

Segundo Mendes (2005), essas anotações compreendem termos úteis para responder perguntas do tipo “onde um determinado produto gênico está localizado na célula?, quais funções ele tem no nível molecular? e a quais processos biológicos estas funções contribuem?” Sendo assim, a GO vem sendo usada no momento em que o biólogo está desenvolvendo sua pesquisa e surgem necessidades de se fazer anotações sobre algumas questões relevantes. Essas anotações são feitas através da terminologia que compõe esta ontologia, a fim de que sejam acessáveis e compreensíveis por outros pesquisadores.

Na figura 2 pode-se observar como essas anotações são inseridas dentro dos bancos de dados genômicos. Geralmente, nesses bancos existem campos específicos para inclusão dos termos da GO pelos pesquisadores. O círculo vermelho mostra os campos onde os termos são inseridos para efeitos de anotações.

A Gene Ontology agrupa os termos em três categorias, referentes a componentes celulares, processos biológicos e funções celulares, que são representados através de suas respectivas definições e relações “is_a” e “part_of”.

Na Categoria Componente Celular⁹ estão incluídos componentes da célula, mas com a condição de ser uma parte de algum objeto maior, quando diz respeito a uma estrutura anatômica (e.x. rough endoplasmic reticulum or nucleus) e componentes de um grupo de produtos genéticos (e.x. ribosome, proteasome or a protein dimer). Esta categoria descreve também os locais, nos níveis das estruturas sub-celulares e dos complexos macromoleculares. Exemplos de componentes celulares incluem “anaphase-promoting complex” e “nuclear membrane”.

⁹ Cf. <http://www.geneontology.org/GO.doc.shtml#cellular_component>.

InterPro: IPR006408 Calcium-translocating P-type ATPase, PMCA-type - Microsoft Internet Explorer

Arquivo Editar Exibir Favoritos Ferramentas Ajuda Links

Simple | Full HTML Version
Click here for help!

InterPro IPR006408 Calcium-translocating P-type ATPase, PMCA-type

Matches	Overview: sorted by AC , sorted by name , of known structure , proteins with splice variants Detailed: sorted by AC , sorted by name , of known structure , proteins with splice variants Table: For all matching proteins , of known structure Architectures
Accession	IPR006408 ATPase-IIB_Ca Matches: 190 proteins
Type	Family
Signatures	Database ID Name Proteins TIGRFAMs TIGR01517 ATPase-IIB_Ca 190
Parent	IPR001757 ATPase, E1-E2 type Tree
Contains	IPR004014 Cation transporting ATPase, N-terminal IPR005834 Haloacid dehalogenase-like hydrolase IPR006068 Cation transporting ATPase, C-terminal IPR008250 E1-E2 ATPase-associated region
Process	GO:0006816 calcium ion transport
Function	GO:0005388 calcium-transporting ATPase activity GO:0005509 calcium ion binding GO:0005524 ATP binding GO:0015085 calcium ion transporter activity
Component	GO:0016020 membrane
Abstract	This family describes the P-type ATPase responsible for translocating calcium ions across the plasma membrane of eukaryotes [1], out of the cell. In some organisms, this type of pump may also be found in vacuolar membranes [2]. In humans and mice, at least, there are multiple isoforms of the PMCA pump with overlapping but not redundant functions. Accordingly, there are no human diseases linked to PMCA defects, although alterations of PMCA function do elicit physiological effects [3]. The calcium P-type ATPases have been characterized as Type IIB based on a phylogenetic analysis which distinguishes this group from the Type IIA SERCA calcium pump [4].
Database links	Enzyme: 3.6.3.8

Taxonomic coverage

Intranet local

Iniciar MSN Messenger 6 Internet Explorer Qualificação Microsoft PowerPoint ... PT 09:22

Figura 2 – Planilha do banco de dados Interpro¹⁰

Na Categoria Processos Biológicos¹¹ são descritos conceitos que representam uma série de eventos realizados por alguém ou funções moleculares organizadas em linha. (Ex: cellular physiological process or signal transduction; ou exemplos mais específicos: pyrimidine metabolism or alpha-glucoside transport). Algumas vezes, pode ser difícil a distinção entre um processo biológico e uma função molecular, mas segundo a documentação da Gene Ontology, a regra geral é que um processo deve ter mais de um passo distinto.

A Categoria Função Molecular¹² é categoria que reúne conceitos que indicam as capacidades ou os trabalhos que um produto genético realiza. Nela podem estar incluídos ações como “transportar, reunir, unir, reunir coisas, e transformar uma

¹⁰ O Interpro é um dos bancos de dados onde pesquisadores da FIOCRUZ inserem suas anotações sobre a comparação de seqüências.

¹¹ Cf. <http://www.geneontology.org/GO.doc.shtml#biological_process>.

coisa em outra”. Desta forma, descreve atividades catalíticas ou atividades de ligação, no nível molecular, tais como catalytic activity, transporter activity, binding; adenylate cyclase activity, Toll receptor binding. A função Molecular da Gene Ontology representa mais funções relacionadas às atividades do que às entidades (moléculas ou complexos) que apresentam as ações e não especificam onde, quando, nem em qual contexto a ação acontece.

Segundo Smith, Williams e Schulze-kremer (2003) a Gene Ontology, apesar de amplamente utilizada, ainda apresenta alguns problemas, a saber:

- Não está claro os tipos de argumentação que são permissíveis na base da hierarquia da GO;
- O motivo da escolha corrente não tem sido preservada e não poderá ser explicada ou reexaminada por terceiros;
- Nenhum procedimento oferecido pela GO pode ser validado;
- Existem regras insuficientes para determinação de como reconhecer se um dado conceito está ou não presente na GO;
- O uso de uma forma de busca pressupõe que todos os conceitos já tenham uma representação padrão única, o que não é o caso.

Além disso, em outros trabalhos, Smith (2003a; SMITH et al., 2005) destacam um problema adicional a GO, que vem justificar a importância desta pesquisa, ou seja, a inconsistência nas relações entre conceitos. “Tais erros derivam-se de uma falta de atenção a princípios ontológicos” (SMITH, 2003a).

Berardini e outros (2004) apresentam o mesmo problema da relação entre conceitos, porém propõem uma complementação do escopo do Gene Ontology através da criação de novos termos e relações para o projeto TAIR (The Arabidopsis Information Resource), que tem por objetivo central a integração de informações de diferentes fontes de dados, apresentando à comunidade científica uma visão compreensiva de cada gene da planta *Arabidopsis thaliana*.

No capítulo seguinte será apresentada uma revisão de literatura sobre as relações conceituais, para que a partir de uma sistematização dessas relações possa se observar a GO com um padrão preestabelecido.

¹² Cf. <http://www.geneontology.org/GO.doc.shtml#molecular_function>.

5 RELAÇÕES CONCEITUAIS NO ÂMBITO DA CIÊNCIA DA INFORMAÇÃO, DA TERMINOLOGIA E DA CIÊNCIA DA COMPUTAÇÃO

A determinação de relações conceituais bem definidas é útil à elaboração de ontologias, na medida em que estas relações garantem a consistência na “adoção de algoritmos baseados em regras de associação” (KIETZ; MAEDACHE; VOLZ, 2000), a consistência na estrutura terminológica (USCHOLD; KING, 1995; FERNANDEZ; GOMEZ-PÉREZ; JURISTO, 1997), e a elaboração da taxonomia que deve compor a estrutura da terminologia (GUARINO, 1995; SURE; STAAB; STUDER, 2002).

O estabelecimento de relações está ligado a dois momentos da construção da ontologia, isto é, na etapa de tratamento da linguagem natural, com a finalidade de garantir consistência na estrutura terminológica, e na etapa de estabelecimento de inferências com a finalidade de garantir a adoção de algoritmos consistentes para a interpretação pela máquina.

Sheth, Arpinar e Kashyap (2003) afirmam que os relacionamentos são fundamentais para Web Semântica, “pois associam os significados às palavras, aos termos e às entidades”, além disso, “são a ‘chave’ para novas percepções e idéias”. Afirmam também que “a descoberta do conhecimento é antes de tudo a descoberta de novos relacionamentos”, evidenciando assim, a importância da utilização de relações nas ontologias para a visão da Web Semântica.

No que diz respeito ao uso de relações para a composição da estrutura terminológica das ontologias, pode-se justificar sua utilidade através da afirmação de Motta (1987), em seu livro “Método relacional”¹³ como nova abordagem para

¹³ “Método analítico que pretende ser instrumento para o estabelecimento de relações em tesauro de forma mais objetiva [...] Consiste na análise das definições dos conceitos que integrarão o sistema, de

construção de tesouros”, que diz que o estabelecimento de relações conceituais pode ser útil em circunstâncias como: “determinação de sinônimos, atualização do sistema, inclusão de termos em uma só categoria e mapeamento de áreas de assunto” (MOTTA, 1987, p. 61). Apesar desta afirmação estar relacionada à construção de tesouros, pode ser estendida também à construção de ontologias.

A ausência de uma base teórica para estabelecimento de relações conceituais justifica a presente pesquisa.

Neste capítulo será apresentada uma revisão de literatura, abrangendo as áreas de Ciência da Informação, Terminologia e Ciência da Computação, apontando as principais contribuições sobre o assunto.

5.1 RELAÇÕES CONCEITUAIS NA CIÊNCIA DA INFORMAÇÃO

Na área de Ciência da Informação destacam-se como autores contribuidores para esta pesquisa, Ranganathan (1963), Neelameghan (1975), Dalhberg (1978a), Lancaster (1986) e Aitchison (1987). Apesar da existência de outras pesquisas sobre o assunto, estes autores foram destacados por serem teóricos que propõem formas inovadoras de sistematização de relações.

É importante destacar aqui que, na área de Ciência da Informação, os estudos não se dão no âmbito do desenvolvimento de Ontologias e sim no âmbito da construção de linguagens documentárias verbais e notacionais. Acredita-se, porém, que possam também ser aplicáveis às ontologias, já que estas, assim como outros tipos de linguagens de indexação, podem ser consideradas como instrumento de representação e recuperação de informação. Além disso, a Ciência da Informação vem apresentando, desde seus primórdios, estudos relativos a sistemas de conceitos e relações conceituais.

Ranganathan (1967), na elaboração de sua Teoria da Classificação Facetada, se limita às relações genéricas e partitivas. Esta limitação se dá devido ao objetivo de seu trabalho, que é o estabelecimento de uma metodologia para classificação e organização de assuntos de livros nas estantes.

Ranganathan apresenta um capítulo em seu “Prolegomena to library Classification” para tratar das relações partitivas, o qual intitula “todo, órgãos e

forma a identificar suas características e na estruturação desses conceitos, tendo em vista as relações existentes entre eles” (MOTTA, 1987, p. 39).

constituintes”. Neste capítulo, ele apresenta as seguintes possibilidades de parte: todo, parte, porção, órgão, constituinte, entidade fluida, sólida e social.

- Todo é uma relação entre um conjunto e um elemento pertencente a esse conjunto. Para Ranganathan essa relação pode ser vista sob dois aspectos: aplicados a um Universo de entidades, todas elas tomadas juntas (conjunto, ex: Floresta) ou aplicado a uma entidade típica de um Universo, a entidade completa (objeto, exemplo: árvore).
- Parte é uma relação entre uma entidade e não o todo dela. Para Ranganathan, essa relação também pode ser vista sob dois aspectos: primeiro, aplicado a algumas, mas não todas as entidades de um Universo de Entidades (exemplo: copa de uma árvore), segundo aplicado a uma entidade típica de um Universo não-todo dele. (exemplo: folha)
- Porção é uma relação entre o todo e uma parte que mantém as características essenciais desse todo. (ex: bolo – fatia do bolo).
- Órgão é uma relação que se dá entre a parte funcional de uma entidade típica de um Universo, ou em outras palavras, é uma relação para designar os diferentes órgãos de um todo que possuem funções diferentes, podendo ser separável do todo, mas se separado, sua função cessa. Ex: pulmão - aparelho respiratório.
- Constituinte é uma relação que pode ser aplicada para uma dada entidade num universo de entidades, em que a última parte de tal entidade não tem qualquer função específica em si mesma, no que diz respeito ao todo, mas que tem sua individualidade, podendo ser parte daquela entidade ou de outra entidade diferente. Ex: bilela-motor de automóvel.
- Entidade fluida, sólida, social são relações que podem ocorrer entre porções, constituintes e órgãos. As relações entre porções ocorre entre entidades fluidas concretas (Ex: Copo de leite -Reservatório de leite. As relações entre constituintes ocorre entre entidades fluídas (Ex: gordura-leite) e sólidas (Ex: tijolo-casa). As relações entre órgão ocorrem entre entidades sólidas (Corpo humano-cabeça) e entidades sociais (Presidência da República- Casa Civil) (GOMES; MOTTA; CAMPOS, 2006).

Outro autor, que nesta perspectiva pode ser mencionado, é Neelameghan (1975), que foi seguidor de Ranganathan, porém trabalhou na perspectiva de construção de Tesouros. Este autor classifica os relacionamentos em:

- Relacionamentos Hierárquicos - que são os de gênero-espécie e as relações de parte-todo.
- Relacionamentos Não-hierárquicos - que são todas as relações que servem para ligar idéias, inclusive as relações de equivalência.

Um fator interessante, porém ambíguo, é que na listagem de relações não hierárquicas, aparecem também relações que Neelameghan considera hierárquicas, como por exemplo, relações de gênero-espécie e as relações parte-todo. Ele faz um estudo a partir de assuntos de livros e verifica os conceitos presentes nesses assuntos e suas relações. Neste estudo serão destacadas as relações que podem ser utilizadas para refletir relações conceituais, em domínios de conhecimento.

Relações
1. Processo e método/dispositivo/mediador usado no processo
2. Processo e produto resultante
3. Processo ocorrendo na seqüência
4. Processo e sua propriedade
5. Processo e propriedade do objeto associado com o processo
6. Processo e pessoa geralmente associada com o processo
7. Propriedade e processo usado com propriedade
8. Entidade e método/dispositivo/mediador usado na produção da entidade
9. Coisa considerada como atributo de outra coisa
10. Coisa e sua aplicação
11. Coisa como material e coisa feita desse material
12. Coisa e sua parte
13. Entidade e sua característica/propriedade
14. Entidade e sua medida ou instrumento de medida
15. Entidade e o lugar onde ela ocorre ou onde é manipulada
16. Entidade e seu predecessor ou precursores
17. Causa e efeito
18. Situação e Condição
19. Quase-sinônimos
20. Duas idéias usadas geralmente, concorrentemente: <ul style="list-style-type: none"> • Contigüidade baseada na definição • Contigüidade baseada no conhecimento empírico

21. Idéias tendo elementos comuns em sua definição
22. Pares de idéias coordenadas, geralmente usadas juntas
23. Duas pessoas interagindo em um contexto especial
24. Idéias aparentemente opostas
25. Idéias antônimas
26. Idéias coordenadas ¹⁴
27. Um termo genérico e outro que deve ser usado mais especificamente
28. Um termo multisignificado que tiver sido limitado em seus significados pelo uso de TG ou TE.
29. Um termo da nota de escopo e outro significado possível para cada termo na linguagem natural
30. Entidade e seu ambiente
31. Entidade 1 e outra entidade 2 associada com a propriedade da primeira entidade
32. Entidade e processo
33. Entidade1 e Entidade2, formando um quadro de considerações da primeira entidade
34. Entidade e o sistema/paradigma/escola de pensamento
35. Entidade estudada em uma relação mútua com outra entidade
36. Entidade e processo desempenhado por ela
37. Processo e o ambiente de sua aplicação
38. Processo e entidade processada
39. Idéias Sinônimas

Quadro 1 – Relacionamentos Não-hierárquicos apresentados por Neelameghan (1978)

Dalhberg, em sua Teoria do Conceito (1978a), enfatiza o papel das características. Para ela, “se dois ou mais conceitos têm pelo menos uma característica em comum, então está claro que deva existir uma relação entre eles”. Segundo Dalhberg (1978b), as relações entre conceitos são divididas em:

¹⁴ Idéias coordenadas são idéias que estão em um mesmo nível ligadas a um conceito superordenado comum.

- Quantitativas - medem a quantidade e a similaridade das características entre conceitos e podem ser de quatro tipos:
 - Identidade Conceitual – as características encontradas em dois conceitos são as mesmas.
 - Inclusão Conceitual – todas as características de um conceito estão contidas em um grande número de características de outro conceito.
 - Interseção Conceitual – as características de dois conceitos se sobrepõem.
 - Disjunção Conceitual – as características de dois conceitos não têm nada em comum.
- Qualitativas – que são subdivididas:
 - Formais ou Categoriais - de acordo com os tipos de conceitos, isto é, de acordo com os referentes conceituais.
 - Materiais-Paradigmáticas ou Ontológicas – de acordo com a categoria fundamental do objeto de um conceito. Esse último pode ser de quatro espécies, a saber:
 - Relacionamento Hierárquico – entre gênero-espécie, e espécie e indivíduos.
 - Relacionamento Partitivo – entre o todo e sua parte, entre duas partes e entre partes e subpartes, a saber: relação partitiva em um sistema natural (ex: as partes do organismo de um animal), relação partitiva em um sistema artificial (ex: as partes de uma máquina), relação partitiva em uma organização humana (ex: país, estado, cidades etc), relação partitiva de um campo de assunto (ex: disciplina e suas partes)
 - Relacionamento de Oposição – relacionamento de oposição contraditória (ex: presença-ausência), de oposição contrária (ex: preto – branco) e PNI, isto é, positivo-neutro-indiferente (ex: favorável – neutro – desfavorável)
 - Funcionais-sintagmáticas¹⁵ - onde se encontram as associativas. Dahlberg (1978b) cita as seguintes categorias de relações funcionais sintagmáticas (instrumentalidade,

¹⁵ Nesta categoria de relações a autora não apresenta exemplos.

causalidade, finalidade, condição, modalidade, potencialidade, co-ocorrência, resultado, lugar e tempo.

Este estudo não abordará as relações quantitativas, pois a intenção do mesmo não é medir o número de características semelhantes ou diferentes, mas identificar as formas de relações conceituais.

Outro autor que aborda a questão das relações entre conceitos na área de Ciência da Informação é Lancaster, em seu livro “Vocabulary control for information retrieval” (1986). Neste livro, Lancaster separa dois capítulos para tratar desse assunto: um para as relações hierárquicas e outro para as relações associativas. Lancaster considera as relações de gênero-espécie e parte-todo como sendo relações hierárquicas, enquanto todas as outras relações são colocadas entre as associativas.

Lancaster (1986) define as relações associativas como sendo “todas aquelas que não podem ser usadas para ligar termos que aparecem em uma mesma hierarquia” e indica uma lista de candidatos para relações associativas.

São elas:

- Entre coisa e sua aplicação.
- Entre um efeito e sua causa..
- Entre uma atividade e um agente.
- Entre a matéria-prima e seu produto.
- Entre duas atividades complementares.
- Entre certas oposições.
- Entre uma atividade e a propriedade associada a ela.
- Entre uma atividade e produto dessa atividade.
- Entre uma coisa e sua parte¹⁶

Aitchison (1987), em seu livro “Construção de Tesouros”, dedica uma parte aos relacionamentos de equivalência, hierárquicos e não hierárquicos.

Esta autora considera como relacionamentos de equivalência os sinônimos e os quase-sinônimos. Como relacionamentos hierárquicos, considera as relações de gênero-espécie, as relações parte-todo e as relações polierárquicas e as relações

¹⁶ A relação partitiva em Lancaster aparece tanto na relação hierárquica, quanto na associativa. (LANCASTER, 1986, p. 40, p. 47)

em renque. Como relacionamentos não-hierárquicos, considera todos os demais termos relacionados.

Aitchison (1987) separa as relações não-hierárquicas em três categorias: coisa, propriedade e processo, ou seja, as relações em renque¹⁷ e repete as relações entre quase-sinônimos.

As relações mencionadas são as seguintes:

- Coisa
 - Coisa/Parte
 - Coisa/Processo
 - Coisa/Propriedade
 - Coisa/Coisa como atributo
 - Coisa/Aplicação
- Propriedade
 - Propriedade/Processo
 - Propriedade/Propriedade como atributo
- Processo
 - Processo/Coisa (agente)
 - Processo/Propriedade
- Quase sinônimo
- Espécies do mesmo gênero

Além desses autores, pode ser utilizada como contribuição a este trabalho a Norma ISO 2788 (1986) de estabelecimento e desenvolvimento de tesauros monolíngües, que divide as relações em: relações de equivalência, relações hierárquicas e relações associativas, que são consideradas pelos autores anteriores como relações não-hierárquicas.

As relações associativas são aquelas que não são relações de equivalência, nem de subordinação, mas que são mentalmente associadas e que também são

¹⁷ O conceito de renques se refere a séries horizontais de conceitos e está em oposição ao conceito de cadeia, que se refere a séries verticais de conceitos. Através dos renques e das cadeias, as relações hierárquicas de gênero-espécie e parte-todo – quando assim consideradas - são reveladas.

expressas em um tesouro. Segundo esta norma, as relações podem ser entre termos de uma mesma categoria e entre termos de categorias diferentes.¹⁸

A norma apresenta, ainda, vários tipos de relações entre categorias, sendo elas:

- Uma disciplina de estudo e objetos e fenômenos por ela estudados.
- Uma operação ou processo e seu agente ou instrumento.
- Uma ação e o produto da ação.
- Uma ação e seu paciente.
- Conceitos relacionados e suas propriedades.
- Conceitos relacionados e suas origens.
- Conceitos ligados por dependência causal.
- Uma coisa e seu agente reverso.
- Um conceito e sua unidade de medida.
- Expressão categorimática¹⁹ e seus nomes incrustados.

Na Ciência da Informação as relações são estabelecidas no âmbito da construção de linguagens documentárias, que se limitam a explicitar relacionamentos categoriais. Para as ontologias são necessárias também relações formais, ou seja, relacionamentos que se explicitam a semântica do que ocorre entre essas categorias.

5.2 RELAÇÕES CONCEITUAIS NA TERMINOLOGIA

Quanto à contribuição da Terminologia, nos baseamos em autores seminais da linha da terminologia normativa, sendo eles, Wüster (1981), Felber (1984) e Sager (1990).

A Teoria Geral da Terminologia (TGT) surgiu como uma disciplina científica, a partir dos trabalhos do engenheiro austríaco Eugen Wüster (1981). Segundo esta teoria, “os termos se definem uns em relação aos outros, formando assim um sistema” (WÜSTER, 1971 apud CAMPOS, 2001b, p. 68). Desta forma, pode-se afirmar que esta teoria coloca, entre outras questões, a relação entre conceitos

¹⁸ A norma ISO 2788 apresenta as relações coordenadas como associativas à falta de um simbolismo mais preciso.

¹⁹ São aquelas expressões em que a categoria vem expressa pelo adjetivo e não pelo substantivo, por exemplo, peixe fóssil é um fóssil.

como centro de suas reflexões: Segundo Campos (2001b, p. 74) “conceitos se relacionam com os outros conceitos em um sistema terminológico, pois são representações mentais das relações que ocorrem entre os objetos na realidade empírica”.

Wüster (1981) separa as relações em Lógicas e Ontológicas. A diferença entre as Relações Lógicas²⁰ e as Ontológicas²¹ é que as primeiras são relações de abstração, ou seja, as relações de gênero e espécie. Já as segundas são relações do ser com o mundo. Para ele, as Relações Ontológicas são aquelas que possuem contigüidade no tempo e no espaço (Relações de Contato) e Relações de Causalidade.

As Relações de Contato são consideradas por ele como a subcategoria mais importante e se auto-explicam a partir de suas espécies que são as Relações de Coordenação e as Relações de Encadeamento.

A principal Relação de Coordenação é a relação de parte-todo. Essa relação pode ocorrer entre o todo e suas partes e entre as próprias partes, sendo considerada uma relação espacial e conseqüentemente uma relação de simultaneidade (WÜSTER, 1981, p. 96). Outras Relações de Coordenação mencionadas por Wüster são a Relação de Inclusão e a Relação de Integração, apesar de evidenciar, através de seu quadro (ver quadro 2) e da simbologia sugerida para representação dessa tipologia de relacionamento, que são tipos de relações distintas. Ele não as define, o que dificulta um pouco o entendimento das mesmas.

Outro tipo de Relação de Contato é a Relação de Encadeamento, que são em conseqüência Relações Temporais e se subdividem em Relação de Antecessão e Relação de Sucessão. Como exemplo desta última, pode-se citar as tabelas cronológicas dos papas ou das dinastias seculares.

As Relações de Causalidade, por sua vez, são relações de parentesco e são subdividas em: Relações entre Gerações, que devem expressar a relação entre duas gerações diferentes e podem ser ascendentes e descendentes - e Relações entre Estágios, que devem expressar relações entre estágios de evolução de um mesmo e único indivíduo ou de uma substância. Sendo assim, as relações entre estágios podem ser: Relações Filogênicas, Relações Ontogênicas e Relações entre

²⁰ As Relações Lógicas repousam na semelhança, ou seja, no fato de que dois conceitos têm, pelo menos, uma característica comum.

²¹ Wüster não apresenta exemplos para suas relações.

Substâncias: que são apresentadas no quadro de Wüster, porém não são definidas em seu texto. Ver quadro 2²².

Relações Conceituais										
Sistema de conceitos (ordenação de conceitos)										
Relações lógicas (relação de abstração, relação de semelhança)			Relações Ontológicas							
			Relações de Contato (relações de Contigüidade)				Relações de causalidade (em particular, relações de parentesco)			
			Relações de Coordenação (em particular, relações parte-todo)		Relações de encadeamento (em particular, relações de sucessão)		Gerações		Estágios	
							Geral	filogênico	ontogênico	substâncias
1	2	3	4	5	6	7	8	9	10	11
Membros das relações	Genérica	>	Conceito de inclusão	>	Predecessor		Ascendente		Ex: larva do ovo	
	Específica	<	Conceito de parte	<	Sucessor		Descendente			
	Conceito expandido		BT							
	Conceito restrito		NT							
	Conceito associado		RT							
Combinações	Determinação	→								
	Conjunto de Conceitos	∧								
	Disjunção de conceitos	∨	integração	y						

Quadro 2 – Classificação das Relações segundo a TGT (WÜSTER, 1981)

Na mesma direção destes estudos, pode-se citar Felber²³ (1984), na área de Terminologia, que apresenta um estudo de relações baseado em Wüster, porém, as classifica de forma diferente, a saber:

²² Wüster (1981) não apresenta maiores explicações sobre o quadro por ele apresentado.

²³ Algumas relações apresentadas por Felber não possuem exemplo, por isso não eles não foram mencionados.

- Relações Lógicas - assim como em Wüster, essa relação pode ser entre dois conceitos (relação de subordinação, relação de interseção lógica, relação de coordenação lógica e relação diagonal lógica) e entre três conceitos (séries lógicas verticais, séries horizontais lógicas e séries de ligações lógicas)
- Relações Ontológicas – formadas pelas relações partitivas, as relações de sucessão e as relações material-produto.
- Relações de Efeito – formadas pelas relações de causalidade, instrumental e descendente.

Quanto às Relações Ontológicas, Felber as separa em:

- Relação Partitiva
- Relação de Sucessão, que é uma relação de contigüidade no tempo
- Relação Material-Produto

A Relação Partitiva, em Felber, por sua vez, é subdividida em:

- Relação Partitiva entre dois conceitos
- Relação Partitiva entre três ou mais conceitos

A Relação Partitiva entre dois conceitos pode ser dos seguintes tipos:

- Relação de Subordinação Partitiva

Essa relação ocorre se um objeto individual consiste nas mesmas partes que outro objeto individual, porém com uma parte adicional. Ex: Avião – asa do avião

- Relação de Interseção Partitiva

Essa relação ocorre entre dois conceitos individuais em comparação; no que diz respeito a suas partes, possui apenas algumas partes em comum com o todo.

Ex: Biologia – química - bioquímica

- Relação de Coordenação Partitiva

Essa relação existe entre dois objetos individuais que representam partes de um todo comum. Ex: Motor de avião – asa de avião

- Relação Diagonal Partitiva

Essa relação existe se duas partes de um todo comum não são relacionadas nem através de subordinação e nem através de coordenação. Ex: Química – Mecânica.

A Relação Partitiva também pode ocorrer entre três ou mais conceitos, podendo apresentar a seguinte tipologia:

- Série Vertical Partitiva – acontece quando a subordinação de conceitos forma uma série partitiva vertical.
- Série Horizontal Partitiva – acontece quando a coordenação de conceitos forma uma série horizontal.
- Link Partitivo – acontece quando há a ligação de dois ou mais objetos individuais, criando uma nova entidade. Esse processo também é chamado de integração. Felber (1984, p. 63) acrescenta ainda sobre essa relação, que “uma ligação ontológica não combina dois ou mais membros de um conceito, mas liga dois ou mais objetos individuais, os quais pertencem a esse conceito”.

O conceito de série apresentado por Felber (1984) se assemelha bastante ao conceito de renques e cadeias apresentado por Ranganathan na Ciência da Informação. Neste sentido, a série vertical estaria para o conceito de cadeia, assim como a série horizontal estaria para o conceito de renque.

Quanto à Relação de Efeito em Felber, ela é subdividida em:

- Relação de Causalidade, ou seja, relação de causa e efeito.
- Relação de Ferramenta, isto é, relação entre a ferramenta e a ação executada com a ferramenta.
- Relação de Descendência, que pode ser:
 - Relação Genealógica, por exemplo, relação entre pai e filho, que são chamadas por Wüster de filogênicas.
 - Relação Ontogenética, por exemplo, relação entre o ovo e a larva, que são chamadas por Wüster de ontogênicas.

A teoria de Wüster²⁴ (1971) origina os estudos do Comitê de Terminologia da ISO (TC-37); no entanto, a Norma ISO 704 apresenta outra classificação para as relações entre conceitos:

- Relações Hierárquicas - que se subdividem em Genéricas e Partitivas;
- Relações Não-Hierárquicas - também chamadas de relações associativas.

Segundo esta norma, as relações partitivas podem ser:

- Similar na natureza (ex: átomo- molécula)
- Essencial – onde uma ou mais parte pode ser compulsória
- Não essencial - onde a parte é opcional

²⁴ Wüster e seus seguidores constituem a chamada Escola de Viena.

- Múltiplas – (ex: uma página dentro de um livro)
- Variáveis – (ex: uma caneta como parte de um tinteiro)

As relações associativas²⁵ se dão em uma conexão temática que pode ser estabelecida entre conceitos, através do conhecimento empírico. Desta forma, a norma destaca a seguinte tipologia de relações associativas:

- Recipiente-conteúdo
- Atividade-ferramenta
- Passos de um ciclo
- Causa-efeito
- Produtor-produto
- Duração-aparelho de medida;
- Profissão-ferramenta típica
- Objeto-ferramenta associada
- Organização-prédio associado

Outro autor importante é Sager (1990), que em seu livro “A practical Course in Terminological Processing”, divide as relações conceituais em: relacionamentos genéricos – que incluem as relações de gênero e espécie; relacionamentos partitivos, mencionando seis tipos de relações.

As partes são constituintes atômicas do todo, isto é, o todo consiste unicamente de muitas partes idênticas, por exemplo, as unidades de uma escala, os caracteres num conjunto de caracteres.

- 1) As partes são um conjunto finito numerado. Ex: Os 54 cartões de um baralho de cartas.
- 2) O todo consiste em vários grupos de partes numerados e não-numerados. Ex: os valores individuais de um baralho de cartas.
- 3) A parte ou as partes são constituintes opcionais do todo. Ex: o rádio do carro.
- 4) A parte é um constituinte e, algumas vezes, o todo. Ex: uma página de preenchimento é parte de um formulário de preenchimento, se a página de preenchimento é a única do formulário de preenchimento, ela constitui o todo do formulário.
- 5) A parte ou partes são alternativas. Ex: mecanismo de alimentação da fita, que tem o rolo da fita ou o carretel de fita, mas não pode ter ambos.

²⁵ As relações associativas na ISO 704 não são apresentadas com exemplo.

Além das relações genéricas e partitivas, Sager (1990, p. 33) menciona outras duas categorias de relacionamentos: os relacionamentos polivalentes - em que inclui as relações de homonímia, que são relações entre conceitos com a mesma grafia e significados diferentes, sendo assim consideradas um fenômeno da língua; e os relacionamentos complexos, em que inclui todos os outros conceitos que se relacionam de forma diferente das consideradas por ele convencionais (genéricos, partitivos e polivalentes).

Para Sager (1990, p. 34), estes últimos relacionamentos “devem ser considerados igualmente importantes e mais reveladores sobre a natureza dos conceitos”. Sager subdivide as relações complexas de duas formas. Em um dado momento, ele tipifica as relações, em um segundo momento ele identifica as categorias dos conceitos, que podem estar em relação, mas não apresenta a tipologia dessas relações. Este trabalho nomeou as relações apresentadas por ele no primeiro momento de relações complexas formais e as apresentadas no segundo momento de relações complexas categoriais, já que ele não usa uma nomenclatura específica para distingui-las.

Conceito 1	Relação	Conceito 2
Queda	<i>is caused by</i>	explosão nuclear
Papel	<i>is a product of</i>	polpa de madeira
Compressibilidade	<i>is a property of</i>	gás
Petróleo	<i>is a product of</i>	refinação de óleo
Temperatura	<i>is a quantitative measure of</i>	calor
Computador	<i>is a instrument for</i>	processamento de dados
Inseticida	<i>is a counterange of</i>	insetos
Caixa de ferramenta	<i>is a container for</i>	ferramentas

Perfuração de diamante	<i>is a method of</i>	perfuração
Ferro	<i>is a material for</i>	construção de ponte
Mina de carvão	<i>is a place for</i>	mineração de carvão

Quadro 3 - Relações Complexas formais

Sager (1990, p. 35) menciona um segundo nível de relacionamentos, que revelam as categorias de conceitos que se relacionam, a saber.

Relacionamentos	Exemplos
Causa-efeito	explosão-queda
Material-produto	aço-chapa de aço
Material-propriedade	óculos-frágil
Material-estado	ferro-corrosão
Processo-produto	tecer-tecido
Processo-instrumento	incisão-bisturi
Processo-método	armazenamento-congelamento seco
Processo-paciente	tintura-têxtil
Fenômeno-medida	luz-watt
Objeto-agente contrário	veneno-antídoto
Objeto-recipiente	ferramenta-caixa de ferramenta
Objeto-material	ponte-ferro
Objeto-qualidade	petróleo-octano elevado
Objeto-operação	peça de broca - perfuração
Objeto-característica	combustível-fumaça
Objeto-forma	livro-brochura
Atividade-lugar	mineração de carvão-mina

Quadro 4 - Relações Complexas Categorias

As relações sugeridas na área de Terminologia também têm em vista um objetivo diferente, isto é, relacionar termos pertencentes às terminologias, glossários, dicionários e léxicos, de maneira geral. Esses instrumentos visam esclarecer terminologia de áreas específicas, servindo como um instrumento regulamentador das expressões utilizadas em uma determinada área. Sendo assim, a complexidade das relações usadas nesses instrumentos ainda não é revelada com todos os detalhes que elas possuem, já que seu uso/processamento não é realizado pela máquina, mas apenas por humanos. As ontologias requerem maior detalhamento das relações, para que possam propiciar a realização de inferências de forma automática, sendo então importante que suas relações possam revelar a categoria dos elementos que se relacionam e também a forma como essa relação é estabelecida entre essas categorias.

As relações nas Ontologias estão presentes em duas partes: no vocabulário terminológico e no vocabulário formal. A Terminologia e a Ciência da Informação, apesar de não necessitarem ser tão específicas na elaboração de seus instrumentos, possuem sólidas teorias que podem auxiliar na construção da parte do vocabulário terminológico da Ontologia. Já a parte do vocabulário lógico pode receber contribuição dos autores da Ciência da Computação e da área de Terminologia, através dos estudos de Sager, que inovou, trazendo para o âmbito dos estudos das relações na área de Terminologia, as relações complexas divididas nas duas formas.

5.3 RELAÇÕES CONCEITUAIS NA CIÊNCIA DA COMPUTAÇÃO

Na área de Ciência da Computação, o destaque foi dado para autores que estão em evidência na literatura sobre relações conceituais, seja na área da Computação, de Inteligência Artificial ou da Ciência Cognitiva, como Sowa (2000), Guarino (1992,1995, 1998) e Merten (1993).

Merten em 1993 publica um texto, tratando especificamente das relações conceituais que ele divide em Relações Hierárquicas e Relações Coordenadas.

As Relações Hierárquicas podem ser definidas como uma relação entre dois conceitos, em que um - o conceito subordinado – depende conceitualmente de outro. Essas relações se subdividem em: hiperonímia, meronímia (também

chamadas de relações de inclusão), relações funcionais, relações de objeto de estudo e relações de representação.

A Relação de Hiperonímia Lógica também chamada de Caracterização ou Relação de Superordenação, são as relações de gênero e espécie. Este autor as define como sendo a “relação entre um objeto e uma propriedade deste objeto, sendo que o valor desta propriedade permite distinguir dois objetos da mesma classe” e sobre elas, ele declara não haver muitas dúvidas.

Já as Relações de Meronímia/Inclusão, podem ser do tipo: Inclusão de classe/lógica, ou seja, Relações Integradoras e Inclusão ontológica/meronímia, também chamadas de “Relação Parte-todo” e ele especifica alguns tipos de relações dessa natureza, a saber:

1. Coleção-elemento. Ex: Floresta- árvore
2. Massa-porção. Ex: bolo-fatia
3. Objeto-constituente. Ex: casa-janela
4. Objeto-constituente acidental. Ex: Livro-índice
5. Objeto-fragmento. Ex: Carro-volante
6. Objeto-unidades. Ex: Ácido nucléico-codons
7. Processo contínuo–ato Ex: Ciclismo - pedalar
8. Processo descontínuo-fase Ex: Replicação - tradução
9. Zona-região. Ex: Bélgica - Brabant

Além das relações de inclusão, Merten (1993) ainda cita as relações de objeto de estudo, isto é, relação entre uma ciência e seu objeto de estudo e relações de representação, ou seja, relação entre um objeto e sua representação (figura, grafo etc).

As relações funcionais são definidas como “relações de caso semântico”, que têm os processos com os “diferentes intervenientes que seu semantismo pede”. Essas relações que chamamos de relações associativas, para ele são relações “mais interessantes”, porque entram freqüentemente na construção de definições. As seguintes relações foram citadas, dentro dessa categoria.:

1. Auxiliar – que é definida como “fator que favorece a realização de um processo” Ex: tradução – ARN de transferência
2. Agente – que é definida como “aquela (e, as, es) que realizam um processo” Ex: Inoculação-hospedeiro

3. Destinatário - que é citada, mas não é definida.
4. Tempo/duração - que é definida como “caracterização temporal (de duração, de freqüência)
5. Instrumento, de início, de fim...) de um processo: Ex: Inoculação – período de inoculação
6. Instrumento - que é definida como “instrumento que serve para realização de um processo” Ex: Inoculação-injeção de inoculação.
7. Paciente - que é definida como “aquele (a, es, as) que sofre um processo, eventualmente numa série de processos encadeados”.
Ex: transcrição - ARN
8. Resultado - . que é definida como “resultado de um processo” Ex: protossíntese-proteína

As Relações Coordenadas são definidas como relações que ocorrem entre dois conceitos, sem que se possa falar de hierarquia entre eles, ou de dependência de um em relação ao outro. Essas relações se subdividem em:

- Relações espaciais – que são relações que mantêm conceitos x e y, se x estiver no interior de u, ou y englobar x. Segundo o autor essa relação nem sempre se distingue facilmente da relação de meronímia. Um exemplo é DOM-TOM não se encontram na França mas fazem parte dela.
- Relações temporais - relação em que um conceito precede ou sucede o outro no tempo. Ex: transcrição-tradução

Guarino (1998) verticaliza o estudo das relações, com a finalidade de melhorar a composição da taxonomia que deve compor a espinha dorsal de uma ontologia. Assim, “Uma taxonomia é a espinha dorsal de uma ontologia, que pode ser ‘encarnada’ com a adição de atributos e outras relações através de nós (como meronímia ou antonímia)” (1998, p. 1).

Este autor apresenta inicialmente toda a problemática sobre inconsistência nos relacionamentos, isto é, o problema da polierarquia para os sistemas de informação, falando da conseqüente sobrecarga da relação “Is-a”, por contemplar apenas as relações is_a e instance_of, que são praticamente de mesma natureza.

Guarino (1998) apresenta exemplos do que considera como problema gerado pela sobrecarga do “Is_a”, ou seja: confusão de sentidos, redução de sentidos,

super-generalização, links de tipos de função duvidosos e confusão de funções taxonômicas.

Em cima destas questões, propõe quatro teorias, através das quais pretende solucionar o problema, sendo elas: Teoria da Parte, Teoria do Todo ou da Integridade, Teoria da Identidade e Teoria da Dependência. De fato, o autor percebe que o que muitas vezes é considerado como uma relação “is_a”, é na verdade, uma relação “part_of”.

Para falar da tipologia de relacionamentos entre conceitos, Guarino remete a questão ao texto de Winston, Chaffin e Herrmann (1987), que são autores da área de Ciência Cognitiva e que na década de 80 já pensavam sobre como as relações entre conceitos podiam se dar na mente humana.

As relações mencionadas por estes autores são relações meronímicas, isto é, relações entre todo e parte, mas além dessas, esses autores também mencionam uma classe de relações chamadas não-meronímicas. As relações são apresentadas da seguinte forma:

Relações Meronímicas

- Componente-Objeto Integral.
- Membro-Coleção.
- Porção-Massa.
- Matéria-Objeto.
- Característica-atividade.
- Lugar-Área.

Relações Não-meronímicas

- Inclusão Topológica - relações entre conteúdo e uma área, ou uma duração temporal e uma coisa contida nela. Ex: Vinho-isopor, prisioneiro-cela; encontro-manhã.
- Inclusão de Classe – também chamada de Relação de hiponímia, considerada por ele como as relações de gênero e espécie.
- Relação de Atribuição – relação entre objeto e seu atributo. Ex: torre-alta.
- Relação Acessória – Ex: brincos-orelha.
- Relação de Propriedade – relação entre um dono e um objeto, que é de sua propriedade. Ex: milionário-dinheiro, autor-copyright.

Sowa (2000) apresenta o estudo das relações de uma forma diferente de todos os outros autores; ao invés de estabelecer duplas de relacionamentos, ele trabalha com uma relação triádica, onde **a** relacionado a **b** gera uma relação do tipo **c**. Sendo assim, têm-se as seguintes relações²⁶:

- Relação de Acompanhamento - relação *objeto-objeto*, onde um objeto participa com outro objeto em algum processo. Ex: Ronnie saiu com Nancy.
- Relação de Quantidade – relação *característica-quantidade*, onde a quantidade é usada como uma medida de alguma característica. O primeiro argumento de uma relação desse tipo é uma característica, que é geralmente expressa por um nome, assim como comprimento, altura, peso, idade, velocidade, temperatura. Ex: O esqui tem um comprimento de 167 cm. Nesse caso o comprimento é a característica e a quantidade é a medida 167cm.
- Relação de Argumento - relação *função-dado*, onde o dado entra na regra para fazer funcionar. Se a função tem mais que uma entrada, os argumentos podem ser distinguidos como Arg1, Arg2 e assim por diante. Essa relação é usada para representação de expressões matemáticas. Exemplo: Raiz (16)= 4.
- Relação de Atributo – relação *objeto-entidade*, onde a relação é entre uma entidade e uma propriedade de algum objeto. Ex. A rosa é vermelha.
- Relação de Base – relação *atributo-tipo*, onde um tipo de regra determina como um atributo está relacionado ao seu objeto. Ex: Sam é um bom músico.
- Relação de Porque - relação *situação-situação*, onde uma situação é causa de uma outra situação e uma causa é correlativa a um efeito. Ex: Você está molhado porque está chovendo.

²⁶ Sowa não distingue claramente o conceito de atributo do de característica, o que gera dificuldade para interpretação de algumas relações. Por exemplo, as relações de quantidade e medida e também as relações de atributo e características, as quais, possuem, inclusive, as mesmas frases como exemplificação.

- Relação de Filho – relação *ser humano-ser humano*. Um ser humano, sendo que um é filho de outro ser humano. Um filho é correlativo para um pai. Ex: Lílian é mãe de Katie.
- Relação de Comparação – relação *atributo-objeto*, onde um objeto serve como um padrão de comparação para algum atributo. Um objeto sendo comparado é correlativo a outro objeto. Ex: Bob é maior que Mary.
- Relação de Característica – relação *entidade-entidade*, onde um tipo de instância é propriedade de entidade. Em inglês, características são geralmente expressas por nomes, tais como forma, cor, comprimento e peso. Ex: A cor da rosa é vermelha.
- Relação de Função – relação *entidade-entidade*, esta é um tipo de relação geral usada para definir todas as regras. Relação entre um objeto (entidade) e o papel (entidade) que ele exerce. Ex: o fiscal tem o papel/função de fiscalizar o cumprimento da lei.²⁷
- Relação de Maneira – relação *processo-entidade*, onde a relação é entre a entidade e a propriedade da entidade relacionada ao processo. Em inglês, maneiras são geralmente expressas por advérbios, tais como: depressa, corajosamente e provisoriamente; sendo assim a relação é entre o processo e a propriedade (entidade). Ex: A ambulância chegou rapidamente.
- Relação de Medida – relação *atributo-quantidade*, onde uma quantidade é usada como medida de alguns atributos. O primeiro argumento da relação de medida está no atributo, o qual é geralmente expresso por um adjetivo, tais como longo, alto, pesado, velho, rápido, ou quente. A relação de medida liga um atributo a uma medida e relação de quantidade liga a característica correspondente à mesma medida. Ex: O Esqui tem comprimento de 167 cm.

²⁷ Esse exemplo foi criado por este trabalho para melhor compreensão da relação.

- Relação de Parte – relação *objeto-objeto*, onde um objeto é um componente de algum outro objeto e como atributo, uma parte é capaz de existir independentemente. Exemplo: Um dedo é parte da mão.
- Relação de Posse - relação ser *vivo-entidade*, onde uma posse está relacionada ao seu possuidor. Ex: O relógio de Niurka parou.
- Relação de Sucessor – relação *ocorrente-ocorrente*, onde um ocorrente ocorre depois de algum outro ocorrente, ou um sucessor é correlativo para seu predecessor. Ex: Depois que Billy comeu o salgadinho, bebeu uma cerveja.

Outros autores na Ciência da Computação, embora não tratando especificamente de relações entre conceitos, citam, em alguma parte de seus trabalhos, relações que podem ser consideradas importantes para a elaboração de Ontologias.

Mahesh (1996) apresenta as relações e as propriedades como sendo tipos de *slots*. Para este autor, um “*slot* é um mecanismo básico de representação de relacionamentos entre conceitos” (MAHESHI, 1996, p. 17), muito dos conteúdos dos *slots* são propriedades, nas quais eles são definidos como conceitos. Desta forma, para ele existem dois padrões: um padrão que se refere aos atributos e outro que se refere às relações.

Um outro fator importante sobre a metodologia apresentada por este autor é essa distinção entre relação e propriedade, muitas das vezes não mencionada pelos autores que tratam da questão. Desta forma, um *slot* para ser considerado uma relação, deve acontecer entre duplas de conceitos recíprocos, sendo a relação, a conexão entre esses dois conceitos; já o *slot*, que visa a representar o padrão “propriedade”, não requer duplas de conceitos recíprocos, pois as propriedades serão afirmativas sobre a classe.

Lagoze e Hunter²⁸ (2001) também se referem às relações como propriedades, mas não distinguem as primeiras das segundas. Sendo assim, apresentam 20 tipos de *relações*, que são, segundo eles, *propriedades*, ou como também chamados na literatura de Ciência da Computação, *slots*, a saber:

²⁸ Estes autores não apresentam exemplos de suas relações.

- Proceeds
- Follows
- isPartOf
- contains
- isSubeventOf
- inContext
- phaseOf
- hasRealization
- hasCopy
- involves
- hasPatient
- usesTool
- hasResult
- destroys
- Creates
- hasAction
- hasPresence
- hasParticipant
- atTime
- inPlace

Slattery (1997) propõe uma classificação para as relações que devem compor uma Ontologia, isto é:

- Relações de Classificação – inclui relações de inclusão de classes.
- Relações Meronímicas – usadas para descrever relacionamentos partitivos.
- Relações Temporais – define o tempo de intervalo e aspectos pontuais de tempo na ontologia.
- Relações Espaciais – usadas para descrever objetos e relacionamentos de espaço entre objetos e figuras;
- Relações de Influência – expressam impacto ou efeito de objeto sobre o outro;
- Relações de Dependência – descrevem um objeto em dependência do outro.

- Relações de Caso - provêm capacidade para expressar o conhecimento estruturado, sem o qual a relação não pode existir.

Apesar de apresentar uma categorização para as relações, este autor não exemplifica os tipos de relações que estão dentro de cada categoria, mas propõe a criação de uma biblioteca de relações, ou seja, “um repositório rico de conhecimento, que inclui relações geralmente usadas e conjunto de definições e caracterização delas para promover reutilização” (SLATTERY, 1997).

Sobre o que foi visto nesta área, pode-se afirmar que é de senso comum para seus autores, que o detalhamento das relações é importante para elaboração de Ontologias consistentes, já que elas podem auxiliar a máquina no processo de inferência de significados. Apesar da literatura sobre essa questão não ter chegado a um ponto comum sobre como essas relações devem ser estabelecidas, pode ser observada a necessidade de dois tipos de relações: uma que é estabelecida entre categorias de conceitos como em Merten (1993) e Winston, Chaffin e Herrmann (1987) e outras que devem expressar a forma como é dada esta relação, como em Sowa (2000), Lagoze e Hunter (2001) e Sager (1990), que apesar de não ser da área de Ciência da Computação, consegue entender e propor níveis de relações distintos.

5.4 COMPARAÇÃO ENTRE AS TEORIAS PARA ESTABELECIMENTO DE RELAÇÕES

O nível de especificação das relações entre conceitos pode variar, de acordo com a linguagem que se pretende representar, já que os instrumentos são desenvolvidos, conforme um objetivo específico. Devido a isto, as abordagens dadas ao estabelecimento de relações na Ciência da Informação, na Terminologia e na Computação, são diferentes porque têm em vista o desenvolvimento de instrumentos distintos.

Em todos os instrumentos, a elaboração de relações entre conceitos faz parte de sua essência, isto é, sem as relações, essas linguagens podem ser consideradas ineficientes. Segundo a ISO 704 (2000, p. 5)

Na organização do conceito em um sistema conceitual, é necessário trazer à mente o campo do conhecimento que originou os conceitos e também considerar as expectativas dos usuários e as metas do sistema que

utilizarão, pois seus objetivos específicos podem sempre variar de acordo com a área do conhecimento que se visa representar.

Sendo assim, é preciso ter bem definidos, no momento da elaboração, os objetivos que se quer atingir através do uso de tal linguagem, pois esses objetivos também poderão influenciar no uso de determinadas relações em detrimento de outras. Os objetivos específicos dessas linguagens podem sempre variar de acordo com as expectativas dos usuários e com as metas do sistema que os utilizarão.

Na Ciência da Informação, a ênfase é dada ao desenvolvimento de linguagens documentárias, ou seja, linguagens para padronização dos assuntos dos documentos e sua posterior recuperação. A classificação dada pelos autores desta área é sempre uma divisão entre relações hierárquicas e não-hierárquicas. Esta classificação diferencia as relações que podem ser visualizadas na parte sistemática das linguagens, isto é, as de gênero-espécie e parte-todo em um Tesouro - das relações que não podem ser visualizadas por falta de uma forma de representação, mas que ainda assim existem, como é o caso das relações associativas (não-hierárquicas).

As relações associativas são assim chamadas por estarem no nível da associação mental dos termos. Mesmo sendo associadas no nível mental, os autores da Ciência da Informação defendem a necessidade de um padrão para estabelecimento desse tipo de relações, afim de que sejam minimizados ruídos na comunicação entre os usuários dos instrumentos.

Outros autores, ainda, defendem que o estabelecimento de relações associativas pode variar de acordo com o contexto para o qual a linguagem é construída, não entrando em maiores detalhes no âmbito de suas teorias, como em Dalhberg (1978a), que assume que não se prende ao detalhamento de relações, porque estas podem variar de área para área.

Na Terminologia, a ênfase é dada ao desenvolvimento de terminologias, glossários, léxicos etc, que são instrumentos de padronização dos termos utilizados por especialistas. O objetivo destes instrumentos é melhorar a troca de informação na comunidade científica. O uso de glossários, por exemplo, para expor em que sentido foi dado no uso de determinado termo. Neste caso, a utilização de relações pode ser importante para apontar outros termos relacionados ao conceito.

Diferentemente da Ciência da informação, a Teoria Geral da Terminologia classifica as relações em lógicas e ontológicas. Nestas últimas, se incluem as partitivas e as associativas.

A Teoria da Terminologia foi desenvolvida para construir glossários, léxicos e/ou dicionários especializados. Ela enfatiza o papel das características para a formação dos conceitos, através da especificação das características de um objeto na definição. É no momento da atribuição das características que se observa a necessidade de estabelecimento de relações e é neste momento que surgem os princípios para estabelecer relacionamentos.

Quando os autores se referem às relações lógicas, estão nomeando as relações de gênero-espécie e as relações de instâncias, que são relações que expressam a natureza de seus referentes. Quando se referem às relações ontológicas, estes autores estão se referindo a todos os outros tipos de relações não-lógicas e por isso incluem entre elas, as relações partitivas, além de todas as outras associativas.

Sager (1990), porém, mesmo sendo da área de Terminologia, propõe uma outra classificação para as relações entre conceitos, isto é: Relacionamentos Genéricos, Relacionamentos Partitivos, Relacionamentos Polivalentes e Relacionamentos Complexos.

Os Relacionamentos Genéricos de Sager (1990) correspondem às relações lógicas, assim denominadas pela Terminologia e às relações de gênero-espécie, como são chamadas pela Ciência da Informação. Os Relacionamentos Partitivos de Sager correspondem aos relacionamentos de parte-todo da Ciência da Informação e apenas a um tipo de relação Ontológica, isto é, a Relação de Contato ou Contigüidade, que são divididas em relação de coordenação e relação de encadeamento.

Já os Relacionamentos Polivalentes correspondem às relações polihierárquicas, mencionadas pela ISO 2788 (1986) e por Aitchison (1987) na categoria relações hierárquicas. Wüster (1981) considera essa relação como lógica, ou para ser mais específico, como relação de combinação lógica.

Os Relacionamentos Complexos correspondem a todas as relações associativas na Ciência da Informação e a um tipo de relação Ontológica na Terminologia, representada, por exemplo, pelas relações de causalidade em Wüster. Apesar de aparecerem em grande parte da literatura da Ciência da Informação e da

Terminologia, esse tipo de relação possui uma característica distinta em Sager, que a torna especial para o estudo das relações conceituais em Ontologias, isto é, elas são divididas em tipos de relação (Relações Complexas Categoriais) e relacionamentos (Relações Complexas formais). Essa forma de apresentar “tipos de relação” ainda não havia aparecido na literatura da área de Terminologia e Ciência da Informação, mas já havia sido apresentada na literatura da Ciência da Computação, como por exemplo em Sowa (2000) e pode-se acreditar que esta divisão das relações seja de grande serventia para o processamento das relações conceituais pela máquina. Sendo assim, Sager pode ser considerado um autor importante para esses estudos, pois consegue observar que as relações podem ocorrer em dois níveis.

Na Ciência da Computação, a ênfase é dada ao desenvolvimento de Ontologias. A necessidade de uma linguagem formal para processamento da máquina faz com que as relações nas ontologias necessitem ser explicitadas seja em forma de verbos, ou verbos preposicionados, conjunções e outros. Os autores desta área também vêm chamando essas relações de funcionais.

As relações complexas mencionadas por Sager são um reflexo das relações que aparecem na literatura da Ciência da Computação com o nome de função.

No que diz respeito ao estabelecimento de relações no âmbito de Ontologias considera-se importante a sistematização de dois níveis de relações, a saber: as relações categoriais e as relações formais.

As relações nas Ontologias podem aparecer em dois momentos: na parte terminológica (que seriam os termos com definições em linguagem natural); e na parte formal – expressas através de linguagem própria para entendimento da máquina.

Na parte terminológica, o uso das relações categoriais, como apresentadas na literatura da Ciência da Informação e da Terminologia, seria de grande benefício, pois ligaria as classes de conceitos.

Já na parte formal não bastaria ligar as classes, mas seria preciso também dizer de que forma se dá essa ligação entre as classes, o que pode ser expresso através das relações sugeridas pela Ciência da Computação, que chamamos aqui de relações formais, já que expressam a forma como as categorias se relacionam. Com as relações estabelecidas nos dois níveis de forma consistente, poderão ser construídas definições formais que possibilitarão o processamento pela máquina.

Sendo assim, as três áreas podem contribuir de alguma forma.

A Ciência da Informação, por exemplo, tem um grande histórico no que diz respeito ao uso em instrumentos de padronização terminológica para tratamento e recuperação de informação e mesmo nas linguagens de tratamento, mais primitivas, como por exemplo, as listas de cabeçalhos de assuntos e os vocabulários controlados. Podiam, muitas das vezes, não ter seus termos definidos conceitualmente, mas os usos de relações remissivas sempre estiveram presentes nesses instrumentos.

A introdução das categorias na elaboração dos vocabulários controlados propiciou um meio para se pensar domínios e pode ser útil ao estabelecimento das relações entre categorias.

Já na Terminologia, por ser próprio da sua natureza, o estudo e a elaboração de terminologias como glossários, léxicos etc, pode contribuir para essa pesquisa com sua metodologia para definição de termos, tendo em vista que as relações podem ser identificadas por elementos que compõem estas definições, ou dito de outra forma, pelas próprias características do conceito.

A Ciência da Computação, por sua vez, é a área onde surge o estudo das Ontologias, enquanto instrumento de recuperação de informação. Todo estudo, no que diz respeito ao uso de uma terminologia padrão, que propicie a comunicação entre máquina-máquina, parte desta área, que também apresenta a necessidade do uso de relações que mostrassem não apenas as categorias de elementos, mas que evidenciassem também a maneira como essas relações entre categorias se estabeleciam, sugerindo assim, o uso de ações/verbos como forma de relacionar as classes e elementos.

Devido a este conhecimento, também a Ciência da Computação pôde perceber a necessidade, de algumas vezes as relações precisarem ser trinárias ou até n-árias, enquanto as relações nas outras áreas eram sempre binárias, (SOWA, 2000, p. 503; SLOMAN, 2005, p. 12), devido ao seu objetivo.

No capítulo seguinte, será apresentada a questão das relações entre conceitos na área Biomédica, tendo em vista que essa é uma das áreas que mais tem desenvolvido pesquisas no âmbito das ontologias.

6 RELAÇÕES CONCEITUAIS NA ÁREA DE BIOMÉDICA

A área Biomédica é uma área que já vem utilizando instrumentos de padronização terminológica para a melhoria da comunicação científica, recuperação da informação e interoperabilidade de dados. Devido a isto, pesquisas com o intuito de melhoria desses instrumentos vêm sendo realizadas no sentido de estabelecimento de relações entre conceitos de forma consistente.

Autores como Schulz, Kumar e Bittner (2005) e Smith e outros (2005) vêm, na mesma perspectiva desta pesquisa, apresentar a problemática das relações conceituais nas Ontologias Biomédicas.

Schulz, Kumar e Bittner (2005) identificam dois tipos de relações no âmbito da Open Biological Ontology (OBO), isto é *is_a* e *part_of* e se concentram na especificação de tipos de relações mereológicas, ou seja, as partitivas. Para eles, a não especificação das partitivas pode gerar inconsistências na interoperabilidade entre pessoas e agentes eletrônicos, principalmente por não distinguir as relações que mudam através do tempo. Eles percebem que existe uma variedade de partitivas a serem utilizadas na descrição de organismos biológicos e seus constituintes e distribuem as relações partitivas em quatro categorias básicas, sendo elas:

- Partes entre entidades materiais e não materiais
- Partes de Continuantes²⁹
- Parte e localização espacial
- Partes de ocorrentes (processos)

²⁹ “Continuantes são aquelas entidades que duram ou continuam a existir através do tempo, enquanto sofrem diferentes tipos de mudanças, incluindo mudança no espaço” (SMITH, 2005).

Outro artigo é o de Smith e outros (2005). Este grupo de pesquisa interdisciplinar reúne pesquisadores da área de Computação, Filosofia, Genética, Bioinformática e Biologia, que juntos identificaram que a questão das relações entre os termos, além de problemática, estava sendo deixada de lado em detrimento do estudo das definições. Eles mencionam a existência de três tipos de relações na OBO, a saber: classe-classe, instância-classe, instância-instância.

Estes autores acreditam que o tratamento dado às relações na área médica pode ser melhorado através do estabelecimento de uma metodologia de definição das relações a serem usadas nas ontologias e, que sendo assim, conseqüentemente, poderão evitar erros nas codificações e nas anotações feitas com essas Ontologias. Eles acreditam também que uma Ontologia de relação pode promover interoperabilidade entre ontologias e suprir novos tipos de argumentação automatizada sobre dimensões temporais e espaciais dos fenômenos médicos e biológicos.

Desta forma, chegaram a um grupo de dez relações que, segundo eles, devem aparecer em ontologias biomédicas. Estas relações foram divididas em quatro grupos, sendo eles: Relações Fundamentais, Relações Espaciais, Relações Temporais e Relações de Participação. A seleção destas relações contou com a ajuda de pesquisadores desta área, os quais prepararam uma lista das principais relações, que deveriam conter um vocabulário que expressasse o conhecimento na área biomédica. Algumas outras relações foram incluídas em comum acordo entre especialista em ontologias e em biomedicina. Chegaram assim, a uma listagem preliminar de relações que devem ser estabelecidas entre os termos, com o objetivo de prover uma representação do conhecimento biomédico de forma eficaz:

- Relações Fundamentais
 - Is_a
 - Part_of
- Relações Espaciais
 - Located_in
 - Contained_in
 - Adjacent_to
- Relações Temporais
 - Transformation_of

- Derives_from
- Preceded_by
- Relações de Participação
 - has_participant
 - has_agent

Outra tentativa de estabelecimento de relacionamentos entre conceitos para recuperação de informação na área biomédica é evidenciada pelo grupo que estuda a TAIR (The Thaliana Arabidopsis Information Resource), que é um Vocabulário Controlado, usado na área de Genoma, assim como a Gene Ontology.

O projeto TAIR é membro da “Gene Ontology Consortium”, é composto de produtos genéticos referentes a genes de plantas. Seu objetivo é descrever a anatomia da Arabidopsis e seus estágios de desenvolvimento (<ftp://ftp.arabidopsis.org/home/tair/Ontologies>), mas estão trabalhando também para estender essa Ontologia com a inclusão de outras espécies de plantas.

A diferença da GO para TAIR é que enquanto a primeira só apresenta relação de *is_a* e *part_of*, a segunda vem utilizando uma tipologia de 21 relações para anotação de seus genes, conforme apresentado em Berardini e outros (2004, p. 751), a saber:

1. Has
2. Located in
3. Involved in
4. Function as
5. Expressed in
6. Related to
7. Functions in
8. Is subunit of
9. Constituent of
10. Expressed during
11. Required for
12. Not involved in
13. Regulates
14. Not expressed in

- 15. Is down-regulated by
- 16. Expressed only in
- 17. Not functions as
- 18. Not located in
- 19. Represses
- 20. Expresses only during
- 21. Not required for

O GO Consortium reconhece a importância da especificação dessa tipologia de relacionamentos pelo TAIR e pode ser que futuramente estes relacionamentos venham a ser utilizados como padrão por todas as ontologias do consórcio.

Uma tentativa ainda de estabelecimento de relações entre conceitos no âmbito da área Biomédica é a do Bodenreider e Burgun (2005), que apresentam um estudo onde são examinadas algumas questões relativas à compatibilidade entre ontologias biomédicas. Para isto, fazem uma breve análise a respeito das relações apresentadas nessas ontologias. Destacamos aqui algumas observações importantes sobre as ontologias OpenCyc, Foundational Model of Anatomy e Menelas.

Segundo os autores, o OpenCyc apresenta apenas relações partitivas. Nessa Ontologia, coisa é “coleção de qualquer coisa”. Ela apresenta relações de partição em Conjunto ou Coleção e relações de Função (Fn), que indica local onde a doença ocorre.

Sobre a Foundational Model of Anatomy, afirmam ser esta uma ontologia que formaliza exclusivamente a representação da estrutura do conteúdo anatômico. Além da taxonomia anatômica, as hierarquias são formuladas na relação transitiva part-of e nas relações anatômicas branch-of e tributary of, que “representam os relacionamentos entre estruturas arbóreas como nervos, artérias, veias e vasos linfáticos”. Essa ontologia também estende os relacionamentos para contains, contained in, adjacent.

A Ontologia MENELAS é um projeto voltado para a compreensão da linguagem natural. As relações são classificadas em 5 espécies de ligações entre os objetos: física, por exemplo, liga objetos contáveis (contains, has for dose, constituted of), objetos reais e pseudo-objetos (component-of). A relação partitiva liga qualquer espécie de objeto físico e tem subdivisões “part fragment” e “part

segment”. Tem, ainda uma relação “functional part” para representar pontos de vista funcionais.

O UMLS³⁰ (2006) - Unified Medical Language System é um sistema de linguagens de tratamento de informação desenvolvida pela US National Library of Medicine, com o objetivo de prover acesso integrado a um grande número de fontes biomédicas através da utilização de vocabulários controlados unificados.

Atualmente o UMLS (2006) conta com 60 vocabulários controlados, cada um com uma natureza, um tamanho e um escopo do domínio biomédico diferente. Cada um desses vocabulários tem um propósito distinto, isto é, uns foram criados para o uso em sistemas de recuperação, outros para registros médicos, outros ainda para uso em sistemas de suporte na decisão médica. Os termos desses vocabulários estão distribuídos em 134 tipos de categorias e se relacionam através de uma tipologia de 54 relações semânticas.

O UMLS tem um Metathesaurus e um grande repositório de conceitos intitulado Semantic Network. O Metathesaurus não impõe restrições às fontes que contribuem com suas informações; com isto, “não fornece a espécie de organização que se requer de uma ontologia”. Mas a Semantic Network (SN) “é desenvolvida independentemente e serve como uma ontologia básica, de alto nível, para o domínio biomédico” (BODENREIDER; BURGUN, 2005, p. 8). Na SN as relações associativas são agrupadas em cinco tipos: **físico** (por exemplo, part_of, branch_of, ingredient_of); **espacial** (por exemplo, location_of, adjacent_to), **funcional** (por exemplo, treats, complicates, causes); **temporal** (por exemplo, co-occurs_with, precedes); **conceitual** (por exemplo, evolution_of, diagnoses).

Os quadros a seguir apresentam as categorias utilizadas na sistematização do vocabulário do UMLS (Quadros 5 e 6) e também as relações semânticas (Quadro 7) utilizados por esta linguagem.

Interessante é observar que as relações no UMLS são expressas em geral por verbos que unem entidade-entidade e entidade-evento.

Assim como nas relações apresentadas pela Ciência da Computação, o que pode ser observado na literatura sobre relações na área Biomédicas é que elas também evidenciam a maneira como essas relações entre categorias se estabelecem, sugerindo assim, o uso de ações/verbos como forma de relacionar as classes de elementos.

³⁰ Cf. <<http://www.nlm.nih.gov/research/umls/>>

Entity	[Entity] (continued)
Physical Object	[Physical Object] (continued)
Organism	Substance
Plant	Chemical
Alga	Chemical Viewed Functionally
Fungus	Pharmacologic Substance
Virus	Antibiotic
Rickettsia or Chlamydia	Biomedical or Dental Material
Bacterium	Biologically Active Substance
Archaeon	Neuroreactive Substance or Biogenic Amine
Animal	Hormone
Invertebrate	Enzyme
Vertebrate	Vitamin
Amphibian	Immunologic Factor
Bird	Receptor
Fish	Indicator, Reagent, or Diagnostic Aid
Reptile	Hazardous or Poisonous Substance
Mammal	Chemical Viewed Structurally
Human	Organic Chemical
Anatomical Structure	Nucleic Acid, Nucleoside, or Nucleotide
Embryonic Structure	Organophosphorus Compound
Anatomical Abnormality	Amino Acid, Peptide, or Protein
Congenital Abnormality	Carbohydrate
Acquired Abnormality	Lipid
Fully Formed Anatomical Structure	Steroid
Body Part, Organ, or Organ Component	Eicosanoid
Tissue	Inorganic Chemical
Cell	Element, Ion, or Isotope
Cell Component	Body Substance
Gene or Genome	Food
Manufactured Object	
Medical Device	
Research Device	
Clinical Drug	

Quadro 5 – Categorias do UMLS (parte 1)³¹

³¹ Fonte: <http://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html>.

<p>[Entity] (continued)</p> <ul style="list-style-type: none"> Conceptual Entity <ul style="list-style-type: none"> Idea or Concept <ul style="list-style-type: none"> Temporal Concept Qualitative Concept Quantitative Concept Functional Concept <ul style="list-style-type: none"> Body System Spatial Concept <ul style="list-style-type: none"> Body Space or Junction Body Location or Region Molecular Sequence <ul style="list-style-type: none"> Nucleotide Sequence Amino Acid Sequence Carbohydrate Sequence Geographic Area Finding <ul style="list-style-type: none"> Laboratory or Test Result Sign or Symptom Organism Attribute <ul style="list-style-type: none"> Clinical Attribute Intellectual Product <ul style="list-style-type: none"> Classification Regulation or Law Language Occupation or Discipline <ul style="list-style-type: none"> Biomedical Occupation or Discipline Organization <ul style="list-style-type: none"> Health Care Related Organization Professional Society Self help or Relief Organization Group Attribute Group <ul style="list-style-type: none"> Professional or Occupational Group Population Group Family Group Age Group Patient or Disabled Group 	<p>Event</p> <ul style="list-style-type: none"> Activity <ul style="list-style-type: none"> Behavior <ul style="list-style-type: none"> Social Behavior Individual Behavior Daily or Recreational Activity Occupational Activity <ul style="list-style-type: none"> Health Care Activity <ul style="list-style-type: none"> Laboratory Procedure Diagnostic Procedure Therapeutic or Preventive Procedure Research Activity <ul style="list-style-type: none"> Molecular Biology Research Technique Governmental or Regulatory Activity Educational Activity Machine Activity Phenomenon or Process <ul style="list-style-type: none"> Human caused Phenomenon or Process <ul style="list-style-type: none"> Environmental Effect of Humans Natural Phenomenon or Process <ul style="list-style-type: none"> Biologic Function <ul style="list-style-type: none"> Physiologic Function <ul style="list-style-type: none"> Organism Function <ul style="list-style-type: none"> Mental Process <ul style="list-style-type: none"> Organ or Tissue Function Cell Function <ul style="list-style-type: none"> Molecular Function <ul style="list-style-type: none"> Genetic Function Pathologic Function <ul style="list-style-type: none"> Disease or Syndrome <ul style="list-style-type: none"> Mental or Behavioral Dysfunction Neoplastic Process <ul style="list-style-type: none"> Cell or Molecular Dysfunction <ul style="list-style-type: none"> Experimental Model of Disease Injury or Poisoning
--	---

Quadro 6 – Categorias do UMLS (parte 2)³²

³² Fonte: <http://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html>.

<p>isa</p> <p>associated_with</p> <p>physically_related_to</p> <p>part_of</p> <p>consists_of</p> <p>contains</p> <p>connected_to</p> <p>interconnects</p> <p>branch_of</p> <p>tributary_of</p> <p>ingredient_of</p> <p>spatially_related_to</p> <p>location_of</p> <p>adjacent_to</p> <p>surrounds</p> <p>traverses</p> <p>functionally_related_to</p> <p>affects</p> <p>manages</p> <p>treats</p> <p>disrupts</p> <p>complicates</p> <p>interacts_with</p> <p>prevents</p> <p>brings_about</p> <p>produces</p> <p>causes</p>	<p>[associated_with] (continued)</p> <p>[functionally_related_to] (continued)</p> <p>performs</p> <p>carries_out</p> <p>exhibits</p> <p>practices</p> <p>occurs_in</p> <p>process_of</p> <p>uses</p> <p>manifestation_of</p> <p>indicates</p> <p>result_of</p> <p>temporally_related_to</p> <p>co_occurs_with</p> <p>precedes</p> <p>conceptually_related_to</p> <p>evaluation_of</p> <p>degree_of</p> <p>analyzes</p> <p>assesses_effect_of</p> <p>measurement_of</p> <p>measures</p> <p>diagnoses</p> <p>property_of</p> <p>derivative_of</p> <p>developmental_form_of</p> <p>method_of</p> <p>conceptual_part_of</p> <p>issue_in</p>
---	---

Quadro 7 – Relações do UMLS (parte 3)³³

³³ Fonte: <http://www.nlm.nih.gov/research/umls/META3_current_relations.html>.

7 SISTEMATIZAÇÃO DAS RELAÇÕES

No levantamento da tipologia de relações, pôde-se encontrar na Ciência da Informação um total de 94 relações, na terminologia 59, na Ciência da Computação 71 e na Bioinformática 66 relações. Com isso, obteve-se um total de 290 ocorrências de relações (Apêndice A)

Das 290 relações foi feita uma análise do que os autores entendiam por cada uma, eliminando assim, as relações sinônimas. Com isso, chegou-se a uma nova listagem composta de 126 relações.

Com estas 126 relações foi realizada uma classificação da tipologia de relações, subdividindo-as no que resolvemos chamar de Relações Categroriais e Relações Formais.

Chamamos de **relações categoriais** aquelas relações que revelam duplas de categorias, ex: coisa-processo, material-produto etc. Esse tipo de relação é bastante utilizada nas áreas de Ciência da Informação e Terminologia, mais precisamente no âmbito da elaboração de instrumentos de padronizações terminológicas tradicionais, como tesouros, glossário etc. Em nosso trabalho, chegamos a uma listagem de 67 relações categoriais (Apêndice B).

Chamamos de **relações formais** aquelas relações que revelam o tipo de relação existente entre as duplas de categorias, ex: `caused_by`, `occurs_in` etc. Esse tipo de relação, apesar de ter aparecido também na literatura de terminologia, é mais utilizado em linguagens de padronização terminológica, que visam ser utilizadas também pela máquina, como as ontologias que precisam de suas relações o mais bem especificadas possível, para promoverem inferência pela máquina. No trabalho de levantamento e seleção chegamos a 59 relações formais (Apêndice B).

De forma geral, a literatura de relações evidencia três tipos de relações, ou seja: as relações genéricas, as relações partitivas e as relações funcionais.

7.1 RELAÇÕES GENÉRICAS

As relações genéricas se dão a partir do princípio lógico de abstração e desta forma respondem a pergunta “é tipo de?” Assim, ocorrem sempre no interior de uma mesma categoria. Esse tipo de relação não apresenta grandes problemas, pois não é mutável, ao contrário, se manifestam sempre da mesma forma. A sistematização foi feita da seguinte forma: Como tipo de relações genéricas, encontramos os seguintes casos:

Relações Categoriais Genéricas	Relações Formais Genéricas
Gênero-espécie	Is_a
Espécie-instância	Is_instance_of

Quadro 8 – Sistematização das relações genéricas

7.2 RELAÇÕES PARTITIVAS

As relações partitivas são aquelas que revelam a relação entre o todo e suas partes. Essas relações se manifestam diferentemente de acordo com a natureza do conceito, ou seja, quando o conceito é classificado como uma personalidade, que é material, o seu todo e parte são sempre um objeto concreto. Entretanto, quando ele se manifesta como um processo, é um objeto ocorrente ou uma etapa, um estágio, uma fase, conforme (SOWA, 2000; SMITH, 2005). Além disso, é importante destacar que elas se dão sempre no interior de uma mesma categoria.

Sobre esse tipo de relação pode ser observado também que a literatura já evidenciou as duplas de ocorrências, que estávamos chamando até agora de relações categoriais, já que essa dupla evidenciava categorias de conceitos, porém não apresentou uma sistematização das possibilidades de relações entre essas duplas.

A sistematização das relações partitivas foi realizada da seguinte forma: primeiramente, foram selecionadas na listagem de relações (Apêndice B) todas as possibilidades de relações categoriais partitivas e para cada uma delas foi criada uma inversa, isto é, quando essa inversa ainda não havia sido mencionada na

literatura (coluna 1). Em seguida, foram verificadas na listagem as possibilidades de relações funcionais que poderiam ocorrer entre as duplas de categorias, criando também, sempre que necessário, as inversas das relações funcionais (coluna 2). Chegamos ao quadro abaixo:

Relações Categoriais Partitivas	Relações Formais Partitivas
<ul style="list-style-type: none"> • Objeto-Constituinte • Constituinte-objeto 	Has_constituent Has_holo_constituent
<ul style="list-style-type: none"> • Objeto-unidade • Unidade-Objeto 	Has_unity Has_holo_unity
<ul style="list-style-type: none"> • Coleção-elemento • Elemento-coleção 	Has_element Has_collection
<ul style="list-style-type: none"> • Massa-porção • Porção-massa 	Has_portion Has_mass
<ul style="list-style-type: none"> • Objeto-material • Material-objeto 	Has_material Has_holo_material
<ul style="list-style-type: none"> • Área-lugar • Lugar-área 	Has_locate Has_holo_locate
<ul style="list-style-type: none"> • Zona-região • Região-zona 	Has_locate Has_holo_locate
<ul style="list-style-type: none"> • Caráter-atividade • Atividade-caráter 	Has_stage Has_holo_stage
<ul style="list-style-type: none"> • Processo contínuo-ato • Ato-processo contínuo 	Has_stage Has_holo_stage
<ul style="list-style-type: none"> • Processo descontínuo-fase • Fase-processo descontínuo 	Has_stage Has_holo_stage
<ul style="list-style-type: none"> • Grupo-membro • Membro-grupo 	Has_member Has_holo_stage
<ul style="list-style-type: none"> • Classe-membro • Membro-classe 	Has_member Has_holo_stage
<ul style="list-style-type: none"> • Substância-particular • Particular-substância 	Has_substance Has_holo_substance

<ul style="list-style-type: none"> • Conjunto-subconjunto • Subconjunto-conjunto 	<p>Has_subset</p> <p>Has_set</p>
<ul style="list-style-type: none"> • Conjunto-elemento • Elemento-conjunto 	<p>Has_element</p> <p>Has_set</p>
<ul style="list-style-type: none"> • Todo/sub-parte temporal • Sub_parte temporal - todo 	<p>Has_subpart</p> <p>Has_holo</p>
<ul style="list-style-type: none"> • Ciência-objeto de estudo • Objeto de estudo - ciência 	<p>Has_branch</p> <p>Has_hole_branch</p>

Quadro 9 – Sistematização das relações partitivas

7.3 RELAÇÕES FUNCIONAIS

As relações funcionais são aquelas que revelam a relação de objeto com o mundo ou a função de um objeto em um determinado contexto. Essas relações podem se manifestar entre categorias e intra-categorias.

As relações entre categorias são relações que se dão entre elementos de categorias distintas e as relações intra-categorias são aquelas que ocorrem entre dois elementos da mesma categoria. Por isso, a sistematização dessas relações se deu da seguinte forma distinta, a saber:

Foi realizado através do sistema de arranjo uma combinação das categorias ranganathianas PMEST, (personalidade³⁴, matéria, energia/processo³⁵, espaço e tempo) como ponto de partida para a reunião dessas relações.

Esse arranjo foi realizado com o intuito de chegar a uma classe mais genérica das relações categoriais. Desta forma, chegou-se às seguintes classes de relações:

- Personalidade-Personalidade
- Personalidade – Matéria
- Personalidade-Processo
- Personalidade-Espaço
- Personalidade-Tempo

³⁴ O mesmo que “continuants”

- Matéria-Personalidade
- Matéria-Processo
- Matéria-Matéria
- Matéria-Espaço
- Matéria-Tempo

- Processo-Personalidade
- Processo-Matéria
- Processo-Processo
- Processo-Espaço
- Processo- Tempo

- Espaço-Personalidade
- Espaço-Medida
- Espaço-Processo
- Espaço-Espaço
- Espaço-Tempo

Conforme a classificação foi sendo realizada, sentiu-se a necessidade de inclusão de outras categorias, que eram previstas por Ranganathan como subcategorias de alguma dessas cinco categorias e que neste trabalho ficaria mais claro a separação das mesmas como uma categoria própria, como é o caso da categoria medida e propriedade. Desta forma, chegou-se a mais 14 relações:

- Personalidade-Medida
- Matéria-Medida
- Processo-Medida
- Espaço-Medida
- Tempo-Medida
- Medida-Medida
- Medida-Propriedade
- Propriedade–Personalidade

³⁵ O mesmo que “occurents”

- Propriedade-Medida
- Propriedade-Processo
- Propriedade-Espaço
- Propriedade-Tempo
- Propriedade-Propriedade
- Propriedade-Medida

Essa distribuição serviu para categorizar as relações apresentadas na literatura, porém, como era esperado, algumas categorias não apresentavam facetas, pois as relações identificadas na literatura não esgotam as possibilidades de manifestação das relações, já que estas podem aparecer de forma diferente dependendo da área.

O passo seguinte foi a identificação da possibilidade de relações formais que podiam se dar entre essas categorias. Para execução desta etapa, utilizou-se a listagem de relações formais identificadas na literatura (Apêndice B). Para cada dupla de ocorrência foram selecionadas todas as possibilidades de relações formais, chegando assim, à seguinte sistematização:

Relações Categorias Funcionais	Relações Formais Funcionais
Personalidade – Personalidade	<ul style="list-style-type: none"> • Is a product of • Is an instrument of • Transformations into • Derives from • Derivated of • Interact with • Manifestation of • Analyses • Measures • Creates • Placed in • Located in • Adjacent to

	<ul style="list-style-type: none"> • Surrounds • Traverses
Processo – processo	<ul style="list-style-type: none"> • Method-of • Interacts with • Is subevent of • Is phase of • Is stage of • Brings about • Prevents
Processo – Personalidade	<ul style="list-style-type: none"> • Brings about • Affects • Disrupts • Destroys • Results in • Placed in
Espaço-Espaço	<ul style="list-style-type: none"> • Adjacent to
Processo – Espaço	<ul style="list-style-type: none"> • Occurs in
Medida – Personalidade	Is a measure of
Medida – Processo	Is a measure of
Personalidade – Processo	<ul style="list-style-type: none"> • Used in • Method of • Caused by • Affects • Performs / Carries out • Used for • Result of • Analyzes • Measures • Diagnoses • Has_agent • Has_patient
Personalidade – Propriedade	<ul style="list-style-type: none"> • Has

Propriedade – Personalidade	<ul style="list-style-type: none">• a property of• la counteragent of
Propriedade-processo	<ul style="list-style-type: none">• property of

Quadro 10 – Sistematização das Relações Funcionais

Essa sistematização evidencia as possibilidades de relações formais existentes entre as duplas de ocorrências de relações categoriais.

Conforme pôde ser observado, existem relações que parecem sinônimas, mas que devem ser analisadas num contexto, para ver se elas podem ser consideradas equivalentes ou não. Nesse último caso, deve se criar regras para quando usar cada relação.

8 APLICAÇÃO DO MODELO DE RELAÇÕES NO ÂMBITO DA GENE ONTOLOGY

Este capítulo trata da análise das definições de uma amostra da Gene Ontology, que foi realizada visando à aplicação do modelo de relações sistematizado no capítulo anterior.

A análise das definições de uma amostra da GO visa identificar a presença de relações formais já identificadas na literatura biomédica ou não. Neste caso, novas relações seriam incluídas.

A lista de termos que serve de base (Anexo A) para análise das definições, é o resultado de uma atividade desenvolvida como parte de um experimento realizado no âmbito do Projeto BioWebDB³⁶, na Fundação Instituto Oswaldo Cruz (FIOCRUZ), Rio de Janeiro. Para identificar os termos de maior relevância para o projeto, foi feita uma busca em 12 diferentes bancos de dados, visando identificar fragmentos de códigos genéticos similares nos fragmentos existentes na FIOCRUZ, relativos ao *Trypanosoma vivax*³⁷. Os termos da Gene Ontology, identificados naqueles bancos de dados, foram listados por ordem de frequência e evidenciariam os termos de maior importância para o projeto.

Apesar do tempo decorrido – cerca de dois anos – a lista pode ainda ser válida, tendo em vista que o objetivo é a identificação de relações formais existentes nas definições. Uma vez que existem padrões para as definições que integram o banco de termos da GO, acredita-se que uma amostra baseada em qualquer outro critério forneceria elementos semelhantes para explorar as definições.

A segunda coluna da lista mostra a frequência de ocorrência dos termos nos bancos de dados. Os 267 termos listados correspondem a 1234 ocorrências.

³⁶ Cf. <<http://www.biowebdb.org/garsa>>.

³⁷ Cf. <<http://www.liebertonline.com/doi/abs/10.1089/omi.2005.9.116>>.

Adotando-se 20% dos termos como base para análise, estaremos abrangendo 54 termos que cobrem aproximadamente 70% das ocorrências, desde 140 até 4 citações. Os restantes 213 termos estão na faixa de 1 a 3 citações.

8.1. A AMOSTRA SELECIONADA

Sobre a amostra selecionada, é preciso destacar que os itens 27, 36 e 41 do Anexo I estão com informações truncadas e foram excluídos. Dos 51 termos restantes, o de número 29 foi substituído posteriormente pela GO. A análise inicial se limitou, portanto, a 50 termos (Apêndice C), os quais mantêm a mesma numeração da base inicial, para facilitar uma eventual consulta.

Como se sabe, a Gene Ontology agrupa os termos em Função molecular, Processo biológico e Componente celular. Seguindo esta classificação, os termos analisados representam:

Função molecular: 24 termos (48%)

Processo biológico: 12 termos (24%)

Componente celular: 14 termos (28%)

Na etapa de análise das definições, algumas dificuldades foram encontradas, o que levou à eliminação de alguns termos, a saber:

item 1 – se tornou obsoleto

item 8 - processo que ainda não pode ser definido por ser desconhecido

itens 12 e 14 – apresentam duas definições cada.

item 34 – é um adjetivo.

item 35 – definição circular

As definições devem descrever de maneira clara o conteúdo do conceito; não é admissível duas descrições para um mesmo conceito. O conceito é sempre designado por um substantivo; o adjetivo é um qualificador. A definição circular é inócua.

A análise aqui apresentada foi estudada a partir da classificação das relações formais e categoriais, mas também procurou-se identificar nas definições, algumas tríades de relações, o que será mostrado mais adiante.

8.2 RELAÇÕES CATEGORIAIS

Na amostra pesquisada, apenas as categorias Personalidade e Energia foram identificadas. Os pares a seguir foram também evidenciados no Capítulo 7.

Foram identificadas 37 ocorrências de relações Personalidade-Personalidade (aqui incluídas as relações Parte-Todo):

- 23 pares em Função molecular: itens 2, 3, 6, 9, 13, 14, 17, 18, 20, 21, 25, 30, 31, 32, 33, 39, 40, 42, 43, 51, 52, 53, 54
- 3 pares em Processo biológico: itens 10, 50, 51
- 11 pares em Componente celular: itens 4, 5, 7, 11, 16, 22, 23, 26, 28, 44, 48

A análise permitiu identificar também uma ocorrência de relação Personalidade-Energia

- no conjunto Função molecular: item 49.

Outra relação identificada através de uma ocorrência foi relação Energia-Energia

- no conjunto Processo biológico: item 19.

A análise evidenciou ainda 5 ocorrências de relações Energia-Personalidade

- no conjunto Processo biológico: itens 19, 38, 45, 46, 47.

Duas observações: 1) o item 30 não está definido na GO, mas fazendo uma pesquisa em termos análogos, pôde-se verificar que existe um padrão de definição, então a relação pôde ser inferida; 2) a definição constante do item 19 revela que há ocorrência de 2 pares.

8.3 RELAÇÕES FORMAIS

Na análise das relações formais, utilizaremos a divisão de classificação da Gene Ontology, isto é, função molecular, processo biológico e componente celular. Transcrevem-se aqui, as instruções constantes dos Guidelines da GO para a elaboração das definições.

8.3.1 Função molecular

As definições deste grupo guardam padrões, aqui transcritos:

term string	definition
x binding	Interacting selectively with x, [brief description of x].
[enzyme] activity	Catalysis of the reaction: [reaction catalyzed by enzyme].
x receptor activity	Combining with x to initiate a change in cell activity.
x transporter activity	Enables the directed movement of x into, out of, within or between cells.

Quadro 11 – Padrão de definição da Categoria Função Molecular

Este padrão fornece algumas relações formais como *interacts with*, *reacts with*, *changes or modifies*, *moves into*, *out of*, *within*, *between*.

Analisando as definições da amostra classificadas neste conjunto, identificaram-se as seguintes relações formais:

- transfers to (itens 2, 9, 33)
- reacts to (itens 3, 13, 14, 20, 30, 39, 40, 43, 52, 53)
- interacts with (itens 6, 21, 31, 32, 42, 54)
- moves along (item 17)
- enables movement with/out of/within/between (item18)
- contributes to (item 25)

É importante observar que o item 1, prejudicado na análise, se refere a um termo que teria sido definido de forma incorreta. Há uma recomendação no sentido de alterar as anotações, fazendo substituição para termos mais adequados GO 42626 e GO 43190, que se referem, respectivamente aos grupos função molecular e componente celular. Consultando o termo GO 42626 'ATPase activity, coupled to transmembrane movement of substances', observamos que existe uma preocupação

com a consistência nas anotações, como se depreende do ‘comentário’: “Consider also annotating to the molecular function term 'ATP binding; GO:0005524' ”; Essa é uma instrução que garante consistência no registro dos termos.

Os textos das definições nos itens 49 e 54 não permitem identificar a relação por um não-especialista.

8.3.2 Processo biológico

Na GO, o processo biológico representa uma série de eventos ou de funções moleculares. Esta definição não está clara, porque parece incluir eventos que estariam na categoria Energia (processo)

Para este grupo são propostos padrões de definição,³⁸, apresentados no quadro abaixo:

x development	The process whose specific outcome is the progression of the x over time, from its formation to the mature structure.v
x morphogenesis	The process by which the anatomical structures of x are generated and organized. Morphogenesis pertains to the creation of form.
x formation	The process that gives rise to x. This process pertains to the initial formation of a structure from unspecified parts.
X structural organization	The process that contributes to creating the structural organization of x. This process pertains to the physical shaping of a rudimentary structure.
x maturation	A developmental process, independent of morphogenetic (shape) change, that is required for x to attain its fully functional state. [description of x]

³⁸ Cf. <<http://www.geneontology.org/GO.process.guidelines.shtml>>.

y cell differentiation	The process whereby a relatively unspecialized cell acquires specialized features of a y cell. (N.B. This may be development of y cell type or a set of cells of y cell type. This will involve the change of a cell or set of cells from one cell identity to another.)
y cell fate commitment	The process whereby the developmental fate of a cell becomes restricted such that it will develop into a y cell.
y cell fate specification	The process whereby a cell becomes capable of differentiating autonomously into a y cell in an environment that is neutral with respect to the developmental pathway. Upon specification, the cell fate can be reversed.
y cell fate determination	The process whereby a cell becomes capable of differentiating autonomously into a y cell regardless of its environment; upon determination, the cell fate cannot be reversed.
y cell development	The process aimed at the progression of a y cell over time, from initial commitment of the cell to a specific fate, to the fully functional differentiated cell.
y cell morphogenesis during differentiation	The process by which the structures of a y cell are generated and organized. This process occurs while the initially relatively unspecialized cell is acquiring the specialized features of a y cell.
y cell maturation	A developmental process, independent of morphogenetic (shape) change, that is required for a y cell to attain its fully functional state. [description of y]

Quadro 12 – Padrão de definição da categoria processo biológico

A GO sugere o uso de qualificadores (quadro abaixo) que auxiliam na identificação de relações formais como *occurring in*, *occurring after*, *occurring prior*.

term string	definition
embryonic x morphogenesis	The process, occurring in the embryo, by which the anatomical structures of x are generated and organized. Morphogenesis pertains to the creation of form.
larval x morphogenesis	The process, occurring in the larva, by which the anatomical structures of x are generated and organized. Morphogenesis pertains to the creation of form.
post-embryonic x morphogenesis	The process, occurring after embryonic development [and prior to (e.g. larval) development], by which the anatomical structures of x are generated and organized. Morphogenesis pertains to the creation of form.

Quadro 13 – Uso de qualificadores na GO

As seguintes relações formais foram identificadas neste conjunto:

- progression of (Itens 10 e 50 – o processo está na relação, que liga duas Personalidades)
- regulates (item 19)
- increases (item 19)
- generates (item 38)
- causes/results in (itens 45, 47)
- occurs after (item 46)
- moves with/out of/within/between (item 51)

Sobre essa categoria, uma observação que deve ser feita, é que não foi identificada a relação no item 37. Os itens 45, 46, 47 poderiam ter suas definições avaliadas por um especialista para verificar se cabem as relações formais propostas,

numa tentativa de sintetizar a ação, já que as relações possuem significado aproximado.

8.3.3 Componente celular

A explicação “um produto genético *está localizado* ou *é um sub-componente de componente celular*” nesta categoria, auxilia na identificação de relações formais.

O quadro a seguir fornece os padrões de definição para o conjunto Componente celular:

Term string	definition
organelle envelope	The double lipid bilayer enclosing the organelle and separating its contents from the rest of the cytoplasm; includes the intermembrane space.
organelle membrane	<p>organelle with a single membrane The lipid bilayer surrounding a(n) organelle.</p> <p>organelle with a double membrane Either of the lipid bilayers that enclose the organelle and form the organelle envelope.</p>
organelle inner membrane	The inner, i.e. lumen-facing, lipid bilayer of the organelle envelope.
organelle outer membrane	The outer, i.e. cytoplasm-facing, lipid bilayer of the organelle envelope.
organelle membrane lumen	The region between the inner and outer lipid bilayers of the organelle envelope.

Quadro 14 – Padrão de definição da categoria Componente celular

De um modo geral, os termos do conjunto Componente celular têm tratamento semelhante (ver itens 11, 16, 22, 23, 26, 48 da amostra). Veja as Relações formais encontradas:

- component of (item 4)
- part of X replicates in Y (Local) (item 7)
- component of X occurs in Y (Local) (item 11)
- part outside of (item 16)
- part of (itens 22, 28)
- part around (item 23)
- part encloses (item 26)
- part within (item 44)
- part around (item 48)

Uma observação a ser feita sobre essa categoria é que o item 5 se refere a um componente desconhecido.

Um estudo minucioso deverá refinar as relações partitivas, disciplinando o uso de parte, componente, constituinte e outros, conforme já identificadas na literatura. A relação partitiva deveria estar presente na terminologia, descritas como partitivas. No entanto, um aspecto que merece investigação é a decisão de incluir a relação 'is a' na maioria dos casos. Neste conjunto, as definições, como se apresentam, não permitem inferir a relação parte-todo, sendo assim, a decisão de se utilizar is_a como relação básica para esse fim merece ser revista.

Uma consideração geral que se pode fazer a respeito dos resultados, é que as definições relativas a processos, são naturalmente de natureza lógica e quando este tipo de definição é empregado, a identificação de relações formais não fica evidente. Nas definições relativas aos conjuntos Função e Componentes, as relações formais são mais evidentes, já que são, em sua essência, ontológicas.

8.4 SISTEMATIZAÇÃO DAS RELAÇÕES

A seguir, será apresentada a tríade de relações identificadas através da análise da GO. A sistematização aqui apresentada obedece à mesma organização

do capítulo 7, em que as relações formais são apresentadas no contexto das relações categoriais.

Relação Categorical	Relação Formal
Personalidade-Personalidade	transferred to reacts to interacts with moves along enables movement with/out of/within/between contributes to transfers to progression of moves with/out of/within/between
Processo-Processo	Regulates increases
Processo-Personalidade	<ul style="list-style-type: none"> • generates • causes/produces/brings about/results in Relações partitivas <ul style="list-style-type: none"> • component of • part in (P como Local) • part of P replicates in P/Local • component of P occurs in P/Local • part outside of • part of • part around • part encloses • part within

Quadro 15 – Sistematização das relações da GO

Através da análise das definições no corpus selecionado, pôde se encontrar na Gene Ontology, novos tipos de relações funcionais, que apesar de não revelados em sua estrutura, estavam implícitos nessas definições.

Essa análise evidenciou relações que já haviam sido sugeridas pela literatura como: *interacts_with*, *development-of*, *creates* e *part_of*, mas também encontrou relações que ainda não haviam sido mencionadas. Isso aconteceu porque cada área se manifesta de forma diferente e as relações serão sempre um reflexo do domínio que se está representando. Neste sentido, para cada ontologia a ser construída, poderão aparecer relações inéditas, ainda não mencionadas pela literatura.

As relações funcionais encontradas nas definições da Gene Ontology que não foram encontradas na literatura são as seguintes:

- *Transfers*
- *Reaction-of*
- *Moves along*
- *Moves_into*
- *Moves_out*
- *Moves_within*
- *Moves_between*
- *Locomotion/Modulates*
- *Movement_into*
- *Movement_out of*
- *Movement_within*
- *Occuring_in*
- *Occurring_after*
- *Occurring_prior*

9 CONSIDERAÇÕES FINAIS

O levantamento bibliográfico realizado mostrou que a literatura, apesar de apresentar bastante material a respeito, não havia ainda um trabalho que sistematizasse esses tipos de relações. Com isso, o intuito desse estudo foi a sistematização das relações apresentadas na literatura. Como o objetivo que se tinha era atender com essa sistematização uma área determinada, isto é, a área biomédica, nosso levantamento considerou não apenas as relações levantadas na literatura, mas também as relações identificadas através da análise de uma ontologia de domínio: a Gene Ontology.

A análise da amostra, apesar de ser em pequeno número, revelou várias relações não encontradas na literatura. Neste sentido, pôde-se identificar 14 novas relações. Esse fato, conforme sinalizado, ocorreu porque as relações entre conceitos se manifestaram diferentemente, de acordo com a área representada, não apenas porque os conceitos são diferentes, mas também porque as respostas que se pretende obter com o uso dos vocabulários também são distintos, fazendo com que na elaboração dessas terminologias haja sempre a necessidade de se pensar quais relações são importantes para tal contexto.

Sobre as definições na Gene Ontology, o que pode ser dito é que apesar da GO possuir padrões para definição de seus termos, o resultado encontrado através da análise das definições do corpus mostrou que nem sempre esse padrão era seguido, dificultando a averiguação da identidade do objeto. Observou-se também a utilização de definição circular.

Uma observação importante é que as definições nem sempre eram ontológicas, isto é, funcionais, não apresentando a relação do conceito com o seu contexto, o que pode ser prejudicial, no momento da elaboração de axiomas para a promoção das inferências.

Uma consideração a respeito da sistematização de relações, é que algumas pareceram ser equivalentes, porém, apenas um estudo maior no âmbito de uma área específica poderia revelar efetivamente se essas relações seriam equivalentes ou não, para isso seria necessário o auxílio de um especialista. Tendo em vista a necessidade de estabelecimento de um padrão, caso essa equivalência seja confirmada, é essencial que se faça uma opção de uso da relação e caso contrário, deve-se estabelecer critérios para quando usar cada um dos relacionamentos.

Na Ciência da Informação, as relações se apresentam sempre de uma forma que se revela apenas as categorias; já na Ciência da Computação e na Bioinformática, pôde-se verificar a existência de outros modelos de relações, ou seja, relações formais, que revelavam apenas o tipo de relação existente entre essas duas categorias.

Nas Ontologias, porém, é necessário que haja os dois tipos de relações: as categoriais, já utilizadas na Ciência da Informação e também as formais, evidenciadas por outras áreas, já que o objetivo do uso de relações no âmbito das ontologias é o raciocínio formal. Sendo assim, a sugestão é o uso de uma tríade de relações que incorpore as relações categoriais, que se manifestam em duplas e as relações formais, que ocorrem entre essas duplas.

O método da Ciência da Informação pode ser considerado válido para sistematização das relações em Ontologias de diversos domínios, propiciando o estabelecimento de regras de inferências consistentes e possibilitando uma recuperação de informação mais precisa nas bases que utilizarem as Ontologias como mecanismos de representação e recuperação de informação.

Neste sentido, como contribuições desse estudo, podemos também mencionar uma ampliação do arcabouço teórico da Ciência da Informação, por ser mais uma tentativa de sistematização de relações ontológicas, que poderão ser, também, utilizadas na elaboração de Instrumentos de Organização e Representação de Conhecimento tradicionais, possibilitando assim, uma recuperação mais eficiente nos Sistemas de Recuperação de Informação.

É necessário dizer ainda, que como toda dissertação, a presente pesquisa encontrou, durante seu desenvolvimento, algumas dificuldades, principalmente por ter sido este um trabalho interdisciplinar. A dificuldade com a terminologia, muitas vezes cheias de homônimos e sinônimos, foi um dos obstáculos encontrados. Outra dificuldade foi o acesso a algumas literaturas e principalmente, a seleção de material

sobre um assunto que ainda é muito novo e que pesquisas para sua consolidação ainda estão sendo desenvolvidas. Podemos citar como obstáculo também, a dificuldade em encontrar um Sistema de Recuperação da Informação, utilizando uma Ontologia.

Como possibilidades de estudos futuros, identificadas por esse trabalho, apontamos a necessidade de estudo das bases teóricas da Ciência da Informação para elaboração de taxonomias e definições, que juntamente com a presente pesquisa, poderá propiciar a construção de ontologias de forma consistente.

REFERÊNCIAS

AITCHISON, J. **Thesaurus construction**: a practical manual. 2. ed. London: ASLIB, 1987.

_____; GILCHRIST, A. **Thesaurus construction**: a practical manual. London: Aslib, 1972.

ALMEIDA, M.B. Roteiro para construção de uma ontologia bibliográfica através de ferramenta automatizada. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n. 2, p. 164-179, jul./dez. 2003.

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

ARENS, Y.; HSU, C.-N.; KNOBLOCK, C.A. Query processing in the SIMS information mediator. In: ARPA/ROME LABORATORY KNOWLEDGE-BASED PLANNING AND SCHEDULING INITIATIVE WORKSHOP, 1996, [S.I.]. **Proceedings...** 1996. Disponível em: <<http://citeseer.ist.psu.edu/51138.html>>. Acesso em: 17 out. 2005.

BERARDINI, T. Z. et al. Functional annotation of the arabidopsis genome using controlled vocabularies. **Plant Physiology**, Rockville, v. 135, n. 2, p. 745-755, 2004.

BERNARAS, A.; LARESGOITI, I.; CORERA, J. Building and reusing ontologies for electrical network applications. In: ECAI'96: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1996, Budapest. **Proceedings...** Budapest: [s.n.], 1996. p. 298-302.

BERNERS-LEE, T. **Information management**: a proposal. Genebra, mar. 1989. Disponível em: <<http://www.w3.org/History/1989/proposal.html>>. Acesso em: 20 set. 2005.

_____; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, [S.I.], v. 284, n. 5, p. 28-37, May 2001. Disponível em: <<http://www.scientificamerian.com/2001/0501issue/0501berners-lee.html>>. Acesso em: 15 jun. 2004.

BECK, H.; PINTO, H.S. **Overview of approach, methodologies, standards, and tools for ontologies**. [2002?]. Disponível em: <<http://www.fao.org/agris/aos/Documents/BackgroundPaper.pdf>>. Acesso em: 03 jun. 2005.

BITTENCOURT, G. **Breve história da inteligência artificial**. [20--?]. Disponível em: <<http://www.das.ufsc.br/gia/history/history.html>>. Acesso em: 03 jun. 2005.

BODENREIDER, O.; BURGUN, A. Biomedical ontologies. In: CHEN, H. et al. **Medical informatics**: advances in knowledge management and data mining in

biomedicine. New York: Springer-Verlag, 2005. In press. Disponível em: <http://mor.nhm.nig.gov/pdf/2005-chapter_medont-06.pdf>. Acesso em: 10 jul. 2006.

BOUAUD, J. et al. Acquisition and structuring of an ontology within conceptual graphs. In: ICCS'94: WORKSHOP ON KNOWLEDGE ACQUISITION USING CONCEPTUAL GRAPH THEORY, 1994, Maryland. **Proceedings...** 1994. p.1-25. Disponível em: <<http://citeseer.ist.psu.edu/bouaud94acquisition.html>>. Acesso em: 15 out. 2005.

BOUAUD, J. et al. Methodological principles for structuring an "ontology". In: ICJAI'95: WORKSHOP ON BASIC ONTOLOGICAL ISSUES SHARING, 1995, Montreal. **Proceedings...** 1995. Disponível em: <<http://www-test.biomath.jussieu.fr/~pz/FTPapiers/Bouaud:IJCAIW95.ps>>. Acesso em: 15 out. 2005.

BRICKLEY, D.; HUNTER, J.; LAGOZE, C. **ABC**: a logical model for metadata interoperability. Harmony Project, Working Paper, 1999. Disponível em: <http://www.ilrt.bris.ac.uk/discovery/harmony/docs/abc/abc_draft.html>. Acesso em: 13 out. 2005.

CAILLIAU, R.; CILLIES, J. **How the web was born**: the story of the world wide web. Oxford: Oxford University Press, 2006.

CAMPOS, L. **OntoGuide-4WS**: uma abordagem sistemática de descrição de serviços web para fins de descoberta dinâmica. 2004. 100f. Dissertação (Mestrado em Informática)–Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.

CAMPOS, M.L.A. **A organização de unidades do conhecimento em hiperdocumentos**: um modelo conceitual como um espaço comunicacional para a realização de autoria. 2001. 186 p. Tese (Doutorado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, 2001a.

_____. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001b.

CAMPOS, M.L.M.; CAMPOS, M.L.A.; CAMPOS, L.M. Web semântica e a gestão de conteúdos informacionais. In: MARCONDES, C.H. et al. **Bibliotecas digitais**: saberes e práticas. 2. ed. Brasília, DF: IBICT, 2006.

CHANDRASEKARAN, B.; JONSON, T.R.; BENJAMINS, V.R. Ontologies: what are they? Why do we need them? **IEEE Intelligent Systems**, [S.l.], v. 14, n. 1, p. 20-26, 1999.

CHANDRASEKARAN, B.; JOSEPHON, J. R.; BENJAMIN, V.R. What are ontologies and why do we need them? **IEEE Intelligent Systems and their applications**, [S.l.], v. 14, n. 1, p. 20-26, 1999.

CORAZZON, R. Ontology: a resource guide for philosophers. 2000. Disponível em: <<http://www.formalontology.it>>. Acesso em: 01 jul. 2006.

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

_____. **Ontical structures and universal classification**. Bangalore: Sarada Ranganathan Endowment, 1978b.

DAN, B.; HUNTER, J.; LAGOZE, C. **ABC**: a logical model for metadata interoperability: harmony discussion note. Oct. 1999. Disponível em: <http://www.ilrt.bris.ac.uk/discovery/harmony/docs/abc/abc_draft.html>. Acesso em: 30 abr. 2004.

DEVEDZIC, V. Ontologies: borrowing from software patterns. **Intelligence**, [Stillwater], v. 10, n. 3, p. 14-24, 1999.

DING, Y. Ontology: the enabler for the semantic web: a review of ontologies with the semantic web in view. **Journal of Information Science**, London, v. 27, n. 6 p. 377-384, 2001.

_____; FOO. S. Ontology research and development: part 1: a review of ontology generation. **Journal of Information science**, London, v. 28, n. 2, p. 123-136, 2002a.

_____. Ontology research and development: part 2: a review of ontology mapping and involving. **Journal of Information Science**, London, v. 28, n. 5, p. 375-388, 2002b.

DOERR. M. The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. **AI Magazine**, California, v. 24, p. 75-92, Sept. 2003.

FELBER, H. **Terminology manual**. Paris: UNESCO, 1984.

FERNANDEZ, M.; GOMEZ-PÉREZ, A.; JURISTO, N. From ontological art towards ontological engineering. **AAAI Technical Report**, California, p. 33-40, jun.1997.

FERNANDEZ-LÓPEZ, M. et al. Overview of methodology for building ontologies. In: WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING METHODS: LESSONS LEARNED AND FUTURE TRENDS, 1999, [S.l.]. **Proceedings...** [S.l.: s. n.], 1999. p. 4-13.

FERNEDA, E. **A recuperação da informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. Disponível em: <www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230>. Acesso em: 01 jul 2006.

FIOCRUZ. **Projeto BiowebDB**. 2005. Disponível em: <<http://www.biowebdb.org.br>>. Acesso em: 06 jan. 2005.

FREITAS, F.L.G. Ontologias. In: _____. **Sistemas multiagentes cognitivos para a recuperação, classificação e extração Integradas de informação da web**. 2002. 205 p. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de Santa Catarina, Florianópolis, 2002. p. 48-80.

GILES, T.R. **Dicionário de Filosofia**: termos e filósofos. São Paulo: EPU, 1993.

GO: gene ontology. [c1999-2006]. Disponível em: <<http://www.geneontology.org/>>. Acesso em: 30 abr. 2006.

GOMES, H.E.; MOTTA, D.F.; CAMPOS, M.L. **Revisitando Ranganathan**: a classificação na rede. 2006. Disponível em: <www.conexaorio.com/bit/revisitando/revisitando.htm>. Acesso em: 10 jul. 2006.

GOMEZ-PEREZ. A evaluation taxonomic Knowledge in ontologies and Knowledge bases. In: WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT, 12, 1999, Alberta. **Proceedings...** [S.l: s.n], 1999.

GRUBER, T.R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, [S.l.], v. 5, p. 199-220, 1993.

GRUNINGER, M. Designing and evaluating generic ontologies. In: ECAI96'S: WORKSHOP ON ONTOLOGICAL ENGINEERING, 1996, [S.l.]. **Proceedings...** [S.l.: s.n.], 1996. p. 53-64.

_____; FOX, M.S. Methodology for the design and evaluation of ontologies. In: IJCAI-95: WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWLEDGE SHARING, 1995, Montreal. **Proceedings...** 1995. Disponível em: <citeseer.ist.psu.edu/82232.html>. Acesso em: 15 out. 2005.

GUARINO, N. Concepts, attributes and arbitrary relations. **Data & Knowledge Engineering**, Trier, v. 8, p. 249-261, 1992.

_____. Formal ontology, conceptual analysis and knowledge representation. **International Journal of Human and Computer Studies**, [S.l.], v. 43, n. 5-6, p. 625-640, 1995.

_____. Some ontological principles for designing upper level lexical resources. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVOLUTION, 1., 1998, Granada. **Proceedings...** 1998. Disponível em: <<http://www.loa-cnr.it/Papers/LREC98.pdf>>. Acesso em: 15 out. 2005.

_____. Understanding, building, and using ontologies. **International Journal of Human Computer Studies**, [S.l.], v. 46, p. 293-310, 1997. Disponível em: <<http://www.loa-cnr.it/Papers/vanHeijst.pdf>>. Acesso em: 03 jun. 2005.

_____; GARRARA, M.; GIARETTA, R. An ontology of meta-level categories. In: THE INTERNATIONAL CONFERENCE OF PRINCIPLES OF KNOWLEDGE REPRESENTATION AND REASONING, 4., 1994, Morgan Kaufmann. **Proceedings...** 1994. Disponível em: <<http://www.loa-cnr.it/Papers/KR94.pdf>>. Acesso em: 15 out. 2005.

GUARINO, N.; CARRARA, M.; GIARETTA, P. **An ontology of meta-level categories**. LADSEB-CNR Int. Rep. 6/93, Preliminary version, Nov. 1993.

HAAV, H. M.; LUBI, T.-L. A survey of concept-based information retrieval tools on the web. In: EAST-EUROPEAN CONFERENCE ADBIS*2001, 5., 2001, Vilnius.

Proceedings... Vilnius: Technika, 2001. v. 2, p. 29-41.

HGP: Human Genome Program, U.S. 2003. Disponível em:

<http://www.ornl.gov/sci/techresources/Human_Genome/publicat/primer2001/index.shtml>. Acesso em: 19 jul. 2005.

HJØRLAND, B. Domain analysis in information science: eleven approaches – traditional as well as innovative”. **Journal of Documentation**, [S.l.], v. 58, n. 4, p. 422-62, 2002.

_____. Library and information science: practice, theory, and philosophical basis. **Information Processing & Management**, Trier, n. 36, p. 501-531, 2004.

_____; ALBRECHTSEN, H. Toward a new horizon in information science: Domain analysis. **Journal of the American Society for Information Science**, [S.l.], v. 46, n. 6, p. 400-425, 1995.

HWANG, C.H. Incompletely and imprecisely speaking: using dynamic ontologies for representing and retrieving information. In: KRDB'99: INTERNATIONAL WORKSHOP ON KNOWLEDGE REPRESENTATION MEETS DATABASES, 6., Linkping, jul. 1999. **Proceedings...**1999. Disponível em:

<<http://citeseer.ist.psu.edu/hwang99incompletely.html>>. Acesso em: 17 out. 2005.

IANNELLA, R.; WAUGH, A. **Metadata**: enabling the Internet. 1997. Disponível em: <<http://www.dstc.edu.au/RDU/reports/CAUSE97>>. Acesso em: 15 jun. 2004.

ISO 2788-1986: documentation guidelines for the establishment and development of monolingual thesauri. 2.ed. [S.l.]: ISO, 1986. 32p.

ISO-DIS-704: **principles and methods of terminology**. Geneva: ISO, 2000.

JASPER, R.; USCHOLD, M. A framework for understanding and classifying ontology applications. In: WORKSHOP ON KNOWLEDGE ACQUISITION MODELING AND MANAGEMENT KAW'99', 12., 1999, Stockholm.

Proceedings...1999. Disponível em:< <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/11-uschold.pdf>>. Acesso em: 15 jun. 2004.

JONES, D.; BENCH-CAPON, T.; VISSER, P. Methodologies for ontology development. In: ITI AND KNOWS CONFERENCE OF THE 15TH IFIP WORLD COMPUTER CONGRESS, 1998, London. **Proceedings...** London: Chapman, 1998. p. 62-75. Disponível em: <<http://cweb.inria.fr/Resources/ONTOLOGIES/methodo-for-onto-dev.pdf>>. Acesso em: 03 jun. 2005.

KBSI-IDEF5: ontology description capture method. [199-?]. Disponível em: <<http://www.idef.com/IDEF5.html>>. Acesso em: 17 out. 2005.

KIETZ, J-U; MAEDCHE, A.; VOLZ, R. A method for semi-automatic ontology acquisition from a corporate intranet. In: EKAW'00: WORKSHOP ON ONTOLOGIES AND TEXT, 2000, Juan-Les-Pins. **Proceedings ... 2000**.

LAGOZE, C.; HUNTER, J. The ABC ontology and mode. **Journal of Digital Information**, [S.l.] v. 2, n. 2, 2001. Disponível em : <<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/>>. Acesso em: 17 out. 2005.

LAMBRIX, P. Ontologies in bioinformatics and systems biology. In: DUBITZKY, W.; AZUAJE, F. (Ed.). **Artificial intelligence methods and tools for systems biology**. New York: Springer-Verlag, 2004. p. 129-146.

LANCASTER, F.W. Farradane e os indicadores relacionais. In: _____. **Vocabulary control for information retrieval**. Trad. de M. A. R. Piedade. Washington, DC: Information Resources Press, 1972.

_____. **Indexação e resumos**: teoria e prática. Brasília, DF: Briquet de Lemos, 1993.

_____. **Vocabulary control for information retrieval**. 2. ed. Arlington: Information Resources Pres, 1986.

LENAT, D.B. CYC: a large-scale investment in knowledge infrastructure. **Communications of the ACM**, [S.l.], v. 58, n. 11, p. 33-38, Nov. 1995.

MAHESH, K. Ontology development for machine translation: ideology and methodology. New Mexico State University CRL report MCCS-96-292, 1996.

MENDES, P. N. **Uma abordagem para construção e uso de Ontologias no suporte à integração e análise de dados genômicos**. 2005. 119p. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

MERTEN, P. **Contribuições das relações conceituais à descrição terminológica**. Bélgica: [s.n.], 1993.

METAXIOTIS, K.S.; PSARRAS, J.E.; ASKOUNIS, D. Building ontologies for production scheduling systems: towards a unified methodology. **Information Management & Computer Security**, [S.l.], v. 9, n. 1, p. 44-50, 2001.

MIZOGUCHI, R.; VAN WELKENHUYSEN, J.; IKEDA, M. **Task ontology for reuse of problem solving knowledge towards very large knowledge bases**. Amsterdam: IOS Press, 1995.

MOREIRA, A. Uso de ontologia em sistemas de informação computacionais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 7, n. 1, p. 49 - 60, jan./jun. 2002.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A.P. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. **DataGramZero**: Revista de Ciência da Informação, [S.l.], v. 5, n. 6, 2004.

MOREIRA, A.; OLIVEIRA, A. Contribuição da terminologia na modelagem de sistemas computacionais. **DataGramZero**: Revista de Ciência da Informação, [S.l.], v. 6, n. 5, out. 2005. Disponível em: <http://www.dgz.org.br/out05/Art_01.htm>. Acesso em: 30 mar. 2006.

MORIN, Edgar. **A cabeça bem feita**: repensar a reforma, reformar o pensamento. 8. ed. Rio de Janeiro: Bertrand Brasil, 2003.

MOTTA, D.F. **Método relacional como nova abordagem para a construção de tesouros**. Rio de Janeiro: SENAI/ DN, 1987.

MOURA, A.M.C. **A web semântica**: fundamentos e tecnologias. Rio de Janeiro: IME, [2002].

MUSEN, M.A. Domain ontologies in software engineering: use of Protégé with the EON architecture. **Methods of Information in Medicine**, [S.l.], n. 37, p. 540–550, 1998.

NAVIGLI, R.; VELARDI, P. Learning domain ontologies from document warehouses and dedicated web sites. **Computational Linguistics**, [S.l.], v. 50, n. 2, 2004.

NEELAMEGHAN, M.L. Non-hierarchical associative relationships: their types and computer generation of RT links. In: SEMINAR ON THESAURUS IN INFORMATION SYSTEMS, 1975, Bangalore. **Proceedings...** Bangalore: Documentation research and Training Center, 1975.

NOY, N. F.; MCGUINNESS, D.L. **Ontology development 101**: a guide to creating your first ontology. Stanford: Stanford University, 2001. Disponível em : <www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf>. Acesso em: 15 out. 2005.

OBO: Open Biomedical Ontologies. [200-?]. Disponível em: <<http://obo.sourceforge.net>>. Acesso em: 30 abr. 2006.

PAZZAGLIA, J.-C.R.; EMBURY, S.M. Bottom-up integration of ontologies in a database context. IN: KRDB'98: WORKSHOP ON INNOVATIVE APPLICATION PROGRAMMING AND QUERY INTERFACES, 1998, Seattle. **Proceedings...** 1998. Disponível em: <<http://citeseer.ist.psu.edu/190102.html>>. Acesso em: 15 out. 2005.

PINTO, H.S.; MARTINS, J.P. Ontologies: how can they be built? **Knowledge and Information Systems**, New York, v. 6, n. 4, p. 441 – 464, 2004.

RAJAN, T.N. Related terms in thesauri. In: SEMINAR ON THESAURUS INFORMATION SYSTEM, Bangalore, 1975. **Proceedings...** Bangalore: Documentation Research and Training Center, 1975.

RAMALHO, R. **Web semântica**: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação. Marília: UNESP, 2006.

RANGANATHAN, S.R. **Colon Classification**. Bombay: Asia Publishing House, 1963.

_____. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.

SAGER, J. C. **A practical course in terminology processing**. London: John Benjamins Publishing Company, 1990.

SCHULZ, S.; KUMAR, A.; BITTNER, T. Biomedical ontologies: what part-of is and isn't. **Journal of Biomedical Information**, [S.l.], Dec. 2005.

SHETH, A.; ARPINAR, I. B.; KASHYAP, V. Relationships at the heart of semantic web: modeling, discovering, and exploiting complex semantic relationships. In: NIKRAVESH, M. et al. **Enhancing the power of the Internet studies in fuzziness and soft computing**. [S.l.: s.n.], 2003.

SLATTERY, N.J. **A study of ontology and its uses in information technology: 1 systems**. 1997. Disponível em: <<http://www.mitre.org/support/papers/swee/papers/slattery/>>. Acesso em: 15 out. 2005.

SMITH, B. Formal ontology and information systems. In: _____. **Beyond concepts: ontology as reality representation**. Amsterdam: IOS Press, 2004. p. 73-84.

_____. Mereotopology: a theory of parts and boundaries. **Data Knowledge Engineering**, [S.l.], n. 20, p. 287-303, 1996.

_____. **Ontology and information systems**. 2002. Disponível em: <[ontology.buffalo.edu/ontology\(PIC\).pdf](http://ontology.buffalo.edu/ontology(PIC).pdf)>. Acesso em: 03 jun. 2005.

_____. The logic of biological classification and the foundations of biomedical ontologies. In: INTERNATIONAL CONFERENCE LOGIC METHODOLOGY PHILOSOPHY OF SCIENCE, 10., 2003, Oviedo. **Proceedings...** 2003. Disponível em: <http://ontology.buffalo.edu/bio/logic_of_classes.pdf>. Acesso em: 10 ago 2006.

_____; CEUSTERS, W. Towards industrial-strength philosophy: how analytical ontology can help medical informatics. **Interdisciplinary Science Reviews**, [S.l.], v. 28, n. 2, p. 106-111, 2003.

SMITH, B.; WILLIAMS, J.; SCHULZE-KREMER, S. The ontology of the gene ontology. **Proc AMIA Symp**, [S.l.], p. 609-613, 2003.

SMITH, B. et al. Relations in biomedical ontologies. **Genome Biology**, [S.l.], v. 6, n. 5, 2005.

SMITH, J.M.; SMITH, D.C.P. Database abstractions: aggregation. **Communications of ACM**, [S.l.], v. 20, n. 6, p. 405-412, jun. 1996.

SOERGEL, D. Some remarks on the information retrieval languages, their analysis and comparison. **Information Storage and Retrieval**, [S.I.], n. 3, p. 219-291, 1967.

_____. The rise of ontologies or the re-invention of classification. *Journal of the American Society for Information Science*, [S.I.], v. 50, n. 12, p. 1119-1120, 1999.

SOWA, J. F. **Building, sharing and merging ontologies**: tutorial. 1999. Disponível em: <<http://users.bestweb.net/~sowa/ontology/ontoshar.htm>>. Acesso em: 30 maio 2004.

_____. **Knowledge representation**: logical, philosophical and Computational foundations. Pacific Grove: Brooks/Cole, 2000.

SLOMAN, M. An ontology-enabled service oriented architecture for pervasive Computing Qun Ni. In: ITCC'05: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY: CODING AND COMPUTING, 2005, London. **Proceedings...** London: Imperial College London, 2005. v. II, p. 797-798.

STAR Lab. **Introducing ontologies**. [2005?]. Disponível em: <www.starlab.vub.ac.be/teaching/introducing_ontologies.pdf>. Acesso em: 03 jun. 2005.

STUMME, G.; MAEDCHE, A. FCA-MERGE: bottom-up merging of ontologies. In: IJCAI: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 17., 2001, Seattle. **Proceedings...** Seattle: [s.n.], 2001. p. 225-230.

SURE, Y; STAAB, S; STUDER, R. **Methodology for development and employment of ontology based knowledge management applications**. 2002. Disponível em: <www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2002_sigmod-methodology.pdf>. Acesso em: 15 out. 2005.

SWARTOUT, B. et al. Toward distributed use of large-scale ontologies. In: KAW'96: KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS WORKSHOP, 10., 1996, Banff. **Proceedings...** 1996. Disponível em: <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.html>. Acesso em: 17 out. 2005.

SWARTOUT, W.; TATE, A. Guest editors' introduction: ontologies. **IEEE Intelligent Systems**, [S.I.], v.14, n.1, p.18-19, Jan. 1999.

TAIR: the arabidopsis information resource. 2006. Disponível em: <<http://www.arabidopsis.org>>. Acesso em: 7 jul. 2006.

TELLO, A.L. Ontologías en la web semântica. In: JORNADAS DE INGENIERIA WEB, 01., 2002. Disponível em: <<http://www.informandote.com/jornadas/ngWeb/articulos/jew02.pdf>>. Acesso em: 30 out. 2004.

TRISTÃO, A.M.D.; FACHIN, G.B.; ALARCON, O.E. Sistema de classificação facetada e tesouros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, DF, v. 33, n. 2, p. 161-171, maio/ago. 2004.

UMLS: Unified Medical Language Systems. 2006. Disponível em: <<http://www.nlm.nih.gov/research/umls>> Acesso em: 11 jun. 2006.

USCHOLD, M. Building ontologies: towards unified methodology. In: EXPERTS SYSTEMS 96: THE ANNUAL CONFERENCE OF THE BRITISH COMPUTER SOCIETY SPECIALIST GROUP ON EXPERT SYSTEMS, 16., 1996, Cambridge. **Proceedings...** 1996. Disponível em: <<http://citeseer.ist.psu.edu/context/151818/260175>>. Acesso em: 15 out. 2005.

_____; GRUNINGER, M. Ontologies: principles, methods and applications. **The Knowledge Engineering Review**, [S.l.], v. 11, n. 12, p. 93-136, 1996.

USCHOLD, M.; KING, M. Towards a methodology for building ontologies. In: IJCAI-95: WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWLEDGE SHARING, 1995, Montreal. **Proceedings...** 1995. Disponível em: <<http://citeseer.ist.psu.edu/uschold95toward.html>>. Acesso em: 15 out. 2005.

VAN ELST, L.; ABECKER, A. Domain ontology agents in distributed organizational memories. In: IJCAI: WORKSHOP ON KNOWLEDGE MANAGEMENT AND ORGANIZATIONAL MEMORIES, 2001, Seattle. **Proceedings...** 2001. Disponível em: <<http://www.dfki.uni-kl.de:8000/~docbase/dokana/WWW/D00000557.pdf>>. Acesso em: 30 mar. 2006.

VAN HEIJIST, G.; SCHREIBER, A.T.; WIELINGA, B.J. Using explicit ontologies in KBS development. **International Journal of Human Computer Studies**, [S.l.], v. 46, n. 2-3, p.183-192, feb./mar. 1997.

VICKERY, B.C. Ontologies. **Journal of Information Science**, London, v. 23, n. 4, p. 227-286, 1997.

W3C. **Semantic web**. 2001. Disponível em : <<http://www.w3c.org/2001/ws>>. Acesso em: 10 jul. 2006.

WACHE, H. et al. Ontology-based integration of information - a survey of existing approaches. In: IJCAI-01 WORKSHOP: ONTOLOGIES AND INFORMATION SHARING, 2001. [S.l.]. **Proceedings...**[S.l.: s.n], 2001. p.108-117. Disponível em: <<http://citeseer.ist.psu.edu/451748.html>>. Acesso em: 15 out. 2005.

WEINSTEIN, P.C. **Ontology-based metadada**: transforming the MARC legacy. Pittsburgh: Digital Libraries, 1998. p. 254-263.

WINSTON, M.; CHAFFIN, R.; HERRMANN, D. A taxonomy of part whole relations. **Cognitive Sci.**, [S.l.], v. 11, p. 417-444, 1987.

WU, S.H.; TSAI, T.H.; HSU, W.L. Domain event extraction and representation with domain ontology. In: IJCAI-03 WORKSHOP ON INFORMATION INTEGRATION ON THE WEB, 2003, Acapulco. **Proceedings...** Acapulco: [s.n.], 2003. p. 33-38.

WÜSTER, E. Begriffs-und themaklassification: Unterschiede in ihrem Weses und ihrer. **Anwendung**, [S.l.], v. 22, n. 3, p. 98-104, 1971.

_____. L'étude scientifique qénérale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, L'informatique et les sciences des chose. In: RONDEAU, G. ; FELBER, E. (Org.). **Textes choisis de terminologie**. Québec: GIRSERM, 1981. V.I : fondéments théoriques de la terminologie, p. 57-114.

APÊNDICE A

Listagem de Relações

1. Todo (conjunto) e um elemento pertencente a esse conjunto. ex: Floresta - árvore). (Ranganathan)
2. Entidade e a Parte (não todo) dela. Ex: copa de uma árvore- folha
3. Todo e uma parte (porção) que mantém as características essenciais desse todo. (ex: bolo – fatia do bolo).
4. Todo e diferentes órgãos, podendo ser separável do todo, mas se separado, sua função cessa. Ex: pulmão-aparelho respiratório.
5. Constituinte de um universo de entidades. Ex: bilela-motor de automóvel.
6. Porções ocorrem entre entidades fluidas e concretas (Ex: Copo de leite - Reservatório de leite)..
7. Constituintes ocorrem em entidades fluídas (Ex: gordura-leite) e sólidas (Ex: tijolo-casa).
8. Órgão e entidades sólidas (Corpo humano-cabeça)
9. Órgão e entidades sociais (Presidência da República- Casa civil)
10. Processo e método/dispositivo/mediador usados no processo Neelameghan
11. Processo e produto resultante
12. Processo ocorrendo na seqüência
13. Processo e sua propriedade
14. Processo e propriedade do objeto associado com o processo
15. Processo e pessoa geralmente associada com o processo
16. Propriedade e processo usado com propriedade
17. Entidade e método/dispositivo/mediador usados na produção da entidade
18. Coisa considerada como atributo de outra coisa
19. Coisa e sua aplicação
20. Coisa como material e coisa feita desse material
21. Coisa e sua parte
22. Entidade e sua característica/propriedade
23. Entidade e sua medida ou instrumento de medida
24. Entidade e o lugar onde ela ocorre ou onde é manipulada
25. Entidade e seu predecessor ou precursores
26. Causa e efeito

27. Situação e Condição
28. Quase-sinônimos
29. Duas idéias usadas geralmente concorrentemente
 - 29a. Contigüidade baseada na definição
 - 29b. Contigüidade baseada no conhecimento empírico
30. Idéias tendo elementos comuns em sua definição
31. Pares de idéias coordenadas geralmente usadas juntas
32. Duas pessoas interagindo em um contexto especial
33. Idéias aparentemente opostas
34. Idéias antônimas
35. Idéias coordenadas
36. Um termo genérico e outro que deve ser usado mais especificamente
37. Um termo multissignificativo que tiver sido limitado em seus significados pelo uso de TG ou TE.
38. Um termo da nota de escopo e outro significado possível para cada termo na linguagem natural
39. Entidade e seu ambiente
40. Entidade 1 e outra entidade 2 , associadas com a propriedade da primeira entidade
41. Entidade e processo
42. Entidade1 e Entidade2, formando um quadro de considerações da primeira entidade
43. Entidade e o sistema/paradigma/escola de pensamento
44. Entidade estudada em uma relação mútua com outra entidade
45. Entidade e processo desempenhado por ela
46. Processo e o ambiente de sua aplicação
47. Processo e entidade processada
48. Idéias sinônimas
49. Relação partitiva em um sistema natural (ex: as partes do organismo de um animal) (Dahlberg)
50. Relação partitiva em um sistema artificial (ex: as partes de uma máquina),
51. Relação partitiva em uma organização humana (ex: país, estado, cidades etc),
52. Relação partitiva de um campo de assunto (ex: disciplina e suas partes)

53. Relacionamento de oposição contraditória (ex: presença-ausência),
54. Relação de oposição contrária (ex: preto – branco),
55. Relação de PNI positivo-neutro-indiferente (ex: favorável – neutro – desfavorável)
56. Instrumentalidade,
57. Causalidade,
58. Finalidade,
59. Condição,
60. Modalidade,
61. Potencialidade,
62. Co-ocorrência,
63. Resultado,
64. Lugar
65. Tempo.
66. Entre coisa e sua aplicação e o inverso; (Lancaster)
67. Entre um efeito e sua causa e o inverso;
68. Entre uma coisa e uma propriedade fortemente associada com ela;
69. Entre a matéria-prima e seu produto;
70. Entre duas atividades complementares;
71. Entre certas oposições;
72. Entre uma atividade e a propriedade associada a ela;
73. Entre uma atividade e produto dessa atividade;
74. Entre uma coisa e sua parte³⁹;
75. Coisa Parte (Aitchison)
76. Coisa/Processo
77. Coisa/Propriedade
78. Coisa/Coisa como atributo
79. Coisa/Aplicação
80. Propriedade/Processo
81. Propriedade/Propriedade como atributo
82. Processo/Coisa (agente)
83. Quase-sinônimo

³⁹ A relação partitiva em Lancaster aparece tanto na relação hierárquica quanto na associativa. Cf. Lancaster (1986, p. 40, p. 47).

84. Espécies do mesmo gênero
85. Uma disciplina de estudo e objetos e fenômenos por ela estudados (ISO 2788)
86. Uma operação ou processo e seu agente ou instrumento
87. Uma ação e o produto da ação
88. Uma ação e seu paciente
89. Conceitos relacionados e suas propriedades
90. Conceitos relacionados a suas origens
91. Conceitos ligados por dependência causal
92. Uma coisa e seu agente reversor
93. Um conceito e sua unidade de medida
94. Frases sincategoremáticas e seus nomes incrustados
95. Relações parte-todo. (Wüster)
96. Relação de inclusão
97. Relação de integração
98. Relação de encadeamento, (relações temporais)
99. Relação de antecessão
100. Relação de sucessão, (ex. pode-se citar as tabelas cronológicas dos papas ou das dinastias seculares.)
101. Relações de causalidade (relações de parentesco)
102. Relações entre gerações - que devem expressar a relação entre duas gerações diferentes e podem ser ascendentes e descendentes –
103. Relações entre estágios – que devem expressar relações entre estágios de evolução de um mesmo e único indivíduo ou de uma substância.
104. Relações filogênicas,
105. Relações ontogênicas
106. Relações entre substâncias.
107. Relação partitivas (Felber)
108. Relação de sucessão
109. Relação material-produto.
110. Relação de Efeito
111. Relação de causalidade,
112. Relação instrumental
113. Relação descendente.

114. Relação Genealógica (ex: pai-filho)
115. Relação Ontogenética, (ex: relação entre o ovo e a larva_
116. Similar na natureza (ex: átomo- molécula); (ISO 704)
117. Essencial – onde uma ou mais partes pode ser compulsória;
118. Não essencial - onde a parte é opcional;
119. Múltiplas – (ex: uma página dentro de um livro);
120. Variáveis – (ex: uma caneta como parte de um tinteiro);
121. As partes são um conjunto finito numerado. Ex: Os 54 cartões de um baralho de cartas. (Sager)
122. O todo consiste em vários grupos de partes numerados e não-numerados. Ex: os valores individuais de um baralho de cartas.
123. A parte ou as partes são constituintes opcionais do todo. Ex: o rádio do carro.
124. A parte é um constituinte e, algumas vezes, o todo. Ex: uma página de preenchimento é parte de um formulário de preenchimento, se a página de preenchimento é a única do formulário de preenchimento, ela constitui o todo do formulário.
125. A parte ou partes são alternativas. Ex: mecanismo de alimentação da fita, que tem o rolo da fita ou o carretel de fita, mas não pode ter ambos.
126. Causa-efeito Ex: Explosão-queda
127. Material-produto - Ex: Aço-Chapa de aço
128. Material-propriedade - Ex: óculos-frágil
129. Material-estado - Ex: Ferro-corrosão
130. Processo-produto - Ex: Tecer-tecido
131. Processo-instrumento - Ex: incisão-bisturi
132. Processo-método - Ex: Armazenamento-congelamento seco
133. Processo-paciente - Ex:Tintura-têxtil
134. Fenômeno-medida - Ex: Luz-watt
135. Objeto-agente contrário- Ex:Veneno-antídoto
136. Objeto-recipiente- Ex:Ferramenta-caixa de ferramenta
137. Objeto-material- Ex:Ponte-ferro
138. Objeto-qualidade- Ex:Petróleo-octano elevado
139. Objeto-operação- Ex:Peça de broca – perfuração
140. Objeto-característica- Ex: Combustível-fumaça

141. Objeto-forma- Ex: Livro-brochura
142. Atividade-lugar - Ex: Mineração de carvão-mina
143. Is caused by
144. Is a product of
145. Is a property of
146. Is a product of
147. Is a quantitative measure of
148. Is a instrument for
149. Is a counterange of
150. Is a container for
151. Is a method of
152. Is a material for
153. Is a place for
154. Auxiliar-processo – Ex: tradução-ARN de transferência (merten)
155. Agente-processo- Ex:Inoculação-hospedeiro
156. Tempo-duração- Ex: Inoculação-período de inoculação
157. Instrumento-processo – Ex: inoculação-injeção de inoculação
158. Paciente-processo – Transcrição de um DNA
159. Processo-resultado – Ex: cozinhar-comida
160. Coleção-elemento – Ex: Coleção-objeto
161. Massa-porção – Ex: bolo-fatia
162. Objeto-constituente – Ex: Avião-motor de avião
163. Objeto-constituente acidental Ex: livro-índice
164. Objeto-fragmento Ex: Carro e fragmentos do carro (obtido através do corte com uma serra para metal)
165. Objeto-material – Ex: Solo-matéria orgânica
166. Objeto-unidades – Ex: ácido nucléico-códon
167. Processo Contínuo-ato
168. Processo descontínuo-fase
169. Zona- região - Ex: Brasil-Rio de Janeiro
170. Ciência-objeto de estudo – Ex: virologia – vírus
171. Objeto-Representação –
172. Relação de Acompanhamento - relação *objeto-objeto*. Ex: Ronnie saiu com Nancy.

173. Relação de Quantidade – relação *característica-quantidade*. Ex: O céu tem um comprimento de 167 cm.
174. Relação de Argumento - relação *função-dado*. Exemplo: Raiz (16)= 4.
175. Relação de Atributo – relação *objeto-entidade*. Ex. A rosa é vermelha.
176. Relação de Base – relação *atributo-tipo*. Ex: Sam é um bom músico.
177. Relação de Porque - relação *situação-situação*. Ex: Você está molhado porque está chovendo.
178. Relação de Filho – relação *ser humano-ser humano..* Ex: Lílian é mãe de Katie.
179. Relação de Comparação – relação *atributo-objeto*. Ex: Bob é maior que Mary.
180. Relação de Característica – relação *entidade-entidade*. Ex: A cor da rosa é vermelha.
181. Relação de Função – relação *entidade-entidade*. Relação entre um objeto (entidade) e o papel (entidade) que ele exerce.
182. Relação de Maneira – relação *processo-entidade*. Ex: A ambulância chegou rapidamente.
183. Relação de Medida – relação *atributo-quantidade*. Ex: O Esqui tem 167 cm de comprimento.
184. Relação de Parte – relação *objeto-objeto*. Exemplo: Um dedo é parte da mão.
185. Relação de Posse - relação *ser vivo-entidade*. Ex: O relógio de Niurka parou.
186. Relação de Sucessor – relação *ocorrente-ocorrente*. Ex: Depois que Billy comeu o salgadinho, bebeu uma cerveja.
187. Proceeds
188. Follows
189. IsPartOf
190. Contains
191. IsSubeventOf
192. InContext
193. PhaseOf
194. HasRealization
195. HasCopy

- 196. Involves
- 197. HasPatient
- 198. UsesTool
- 199. HasResult
- 200. Destroys
- 201. Creates
- 202. HasAction
- 203. HasPresence
- 204. HasParticipant
- 205. AtTime
- 206. InPlace
- 207. Relações de Classificação – inclusão de classes;
- 208. Relações Meronímicas –relacionamentos partitivos.
- 209. Relações Temporais – define o tempo de intervalo e aspectos pontuais de tempo;
- 210. Relações Espaciais – usados para descrever objetos e relacionamentos de espaço entre objetos e figuras;
- 211. Relações de Influência – expressam um impacto ou efeito de objeto sobre o outro;
- 212. Relações de dependência – descrevem um objeto em dependência do outro.
- 213. Relações de caso - provêem capacidade para expressar o conhecimento estruturado, sem o qual a relação não pode existir;
- 214. Componente-Objeto Integral (Winston)
- 215. Membro-Coleção
- 216. Porção-Massa
- 217. Matéria-Objeto
- 218. Característica-atividade
- 219. Lugar-Área
- 220. Inclusão Topológica – conteúdo-área, ou duração temporal e uma coisa contida nela.
- 221. Inclusão de Classe – gênero e espécie.
- 222. Relação de Atribuição – relação entre objeto e seu atributo.
- 223. Relação Acessória – Ex: brincos-orelha.

- 224. Relação de Propriedade – dono - propriedade.
- 225. Partes entre entidades materiais e não materiais (kumar)
- 226. Partes de Continuantes
- 227. Parte e localização Espacial
- 228. Partes de Ocorrentes (processos)
- 229. Located_in (SMITH, 2005)
- 230. Contained_in
- 231. Adjacent_to
- 232. Transformation_of
- 233. Derives_from
- 234. Preceded_by
- 235. Has_participant
- 236. Has_agent
- 237. Branch-of (Foundational Model of Anatomy)
- 238. Tributary of
- 239. Contains, (Menelas)
- 240. Has for dose,
- 241. Constituted of;
- 242. Component-of
- 243. Part fragment
- 244. Part segment.
- 245. Functional part
- 246. Consist_of (UMLS part)
- 247. Contains
- 248. Connected_to
- 249. Interconnects
- 250. Branch_of
- 251. Tributary_of
- 252. Ingredient-of
- 253. Location_of (spatially)
- 254. Adjacent_of
- 255. Surrounds
- 256. Traverses
- 257. Affects (functional)

- 258. Manages
- 259. Treats
- 260. Disrupts
- 261. Complicates
- 262. Interact_with
- 263. Brings_about
- 264. Prevents
- 265. Produces
- 266. Causes
- 267. Performs
- 268. Carries_out
- 269. Exhibits
- 270. Practices
- 271. Occurs in
- 272. Process_of
- 273. Uses
- 274. Manifestation_of
- 275. Indicates
- 276. Results_of
- 277. Co-occurs-with
- 278. Precedes
- 279. Conceptually_related_to
- 280. Evaluation_of
- 281. Degree_of
- 282. Analyze_of
- 283. Assessed_of
- 284. Measurement_of
- 285. Measures
- 286. Diagnoses
- 287. Property_of
- 288. Development_of
- 289. Method_of
- 290. Issue_in

APÊNDICE B

Listagem de Relações sem sinônimos

Relações Categoriais

1. Coleção-elemento
2. Massa-porção
3. Objeto-constituente
4. Objeto-constituente acidental
5. Objeto-fragmento
6. Objeto-unidade
7. Ciência-objeto de estudo
8. Objeto-representação
9. Recipiente-conteúdo
10. Causa-efeito
11. Produtor-produto
12. Profissão-ferramenta
13. Produto-instrumento
14. Organização-prédio associado
15. Coisa-coisa considerada como atributo de outra coisa
16. Coisa-aplicação
17. Coisa-propriedade
18. Entidade-medida
19. Situação-condição
20. Idéias quase sinônimas
21. Duas idéias usadas simultaneamente
22. Idéias tendo elementos comuns nas definições
23. Idéias sendo usadas em combinação
24. Idéias aparentemente opostas
25. Expressões categorimáticas-nomes incrustados
26. Idéias antônimas
27. Disciplina –campo de estudo
28. Objeto-estrutura social
29. Agente-instrumento

30. Coisa-origem
31. Contato-agente
32. Objeto-predecessor
33. Pai-filho
34. Ovo-larva
35. Estágio1 da substância-estágio 2 da substância
36. Influência-objeto influenciado
37. Oposição
38. Medida específica- aparelho de medida
39. Pessoa-outra pessoa que interaje no mesmo contexto
40. Objeto-outro objeto que depende dele
41. Objeto-material
42. Matéria-prima produto
43. Material-processo
44. Material-estado
45. Substância-propriedade
46. Auxiliar-processo
47. Agente-Processo
48. Paciente-Processo
49. Processo-resultado
50. Instrumento-processo
51. Processo contínuo-ato
52. Processo descontínuo-fase
53. Processo e processo seqüente
54. Processo-propriedade do processo
55. Processo-propriedade do objeto associado ao processo
56. Propriedade e processo executado por ele
57. Processo-método
58. Processo-material
59. Processo-local
60. Processo-mediador usado no processo
61. Processo-Método
62. Dois processos complementares
63. Zona-região

64. Objeto-localização
65. Tempo-duração
66. Duração-aparelho de medida
67. Intervalo de tempo-aspectos do tempo

Relações Formais

1. Adjacent_to
2. Affects
3. Analyzes
4. Associated_with
5. atTime
6. Branch_of
7. Carries_out / has realization
8. Complicates
9. Consist_of
10. Contains
11. Co-occurs with
12. Creates
13. Degree_de
14. Derivative_de
15. Derives from
16. Destroys
17. Diagnoses
18. Disrupt
19. Exhibits
20. Follow
21. Functionally_related_of
22. Has_action
23. Has_agent
24. Has_copy
25. Has_participant
26. Has_patient
27. In_context

28. Indicates
29. Interacts_with
30. Is a container of/contained_in
31. Is a counteragent Of
32. IS a instrument of / used tool
33. Is a material of
34. Is a place for / inPlace
35. Is a product of
36. Is a property of
37. Is a quantitative measure of
38. Is caused by / brings_about
39. Is_subevent_of
40. Located_in/Location_of
41. Manages
42. Manifestation_of
43. Measure
44. Method_of
45. Occurs_in
46. Part_of
47. Perform
48. Phase_of
49. Preceded_by
50. Precedes
51. Prevents
52. Produces
53. Result_of/ has result
54. Surround
55. Temporally_rellated_to
56. Transformation
57. Travesses
58. Treats
59. Uses

Total: 126 Relações

APÊNDICE C

Definições analisadas*

1 -id: GO:0004009

name: ATP-binding cassette (ABC) transporter activity

namespace: molecular_function

def: "OBSOLETE (was not defined before being made obsolete)." [GOC:jl]

comment: This term was made obsolete because ABC transporters are a protein family rather than a functional grouping. To update annotations, consider the molecular function term 'ATPase activity, coupled to transmembrane movement of substances ; GO:0042626' and its children and the cellular component term 'ATP-binding cassette (ABC) transporter complex ; GO:0043190'.

is_obsolete: true

2 -id: GO:0004672

name: protein kinase activity

namespace: molecular_function

alt_id: GO:0050222

def: "Catalysis of the transfer of a phosphate group, usually from ATP, to a protein substrate." [GOC:jl]

is_a: GO:0016301 ! kinase activity

is_a: GO:0016773 ! phosphotransferase activity, alcohol group as acceptor

3 -id: GO:0004674

name: protein serine/threonine kinase activity

namespace: molecular_function

def: "Catalysis of the reaction: ATP + a protein serine/threonine = ADP + protein serine/threonine phosphate." [GOC:bf]

is_a: GO:0004672 ! protein kinase activity

4 -id: GO:0005737

name: cytoplasm

namespace: cellular_component

def: "All of the contents of a cell excluding the plasma membrane and nucleus, but including other subcellular structures." [ISBN:0198547684]

is_a: GO:0044424 ! intracellular part

5 -id: GO:0008372

name: cellular component unknown

namespace: cellular_component

def: "Used for the annotation of gene products whose localization is not known or cannot be inferred." [GOC:ma]

is_a: GO:0005575 ! cellular_component

6 -id: GO:0005515

* Mantém-se aqui a mesma numeração dos termos listados no **Anexo 1** para facilitar consulta eventual àquela lista.

name: protein binding
 namespace: molecular_function
 def: "Interacting selectively with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules)." [GOC:go_curators]
 is_a: GO:0005488 ! binding

7 -id: GO:0005634
 name: nucleus
 namespace: cellular_component
 def: "A membrane-bounded organelle of eukaryotic cells in which chromosomes are housed and replicated. In most cells, the nucleus contains all of the cell's chromosomes except the organellar chromosomes, and is the site of RNA synthesis and processing. In some species, or in specialized cell types, RNA metabolism or DNA replication may be absent."
 is_a: GO:0043231 ! intracellular membrane-bound organelle

8 -id: GO:0000004
 name: biological process unknown
 namespace: biological_process
 def: "Used for the annotation of gene products whose process is not known or cannot be inferred." [GOC:sgd_curators]
 is_a: GO:0008150 ! biological_process

9 -id: GO:0016301
 name: kinase activity
 namespace: molecular_function
 def: "Catalysis of the transfer of a phosphate group, usually from ATP, to a substrate molecule." [ISBN:0198506732]
 is_a: GO:0016772 ! transferase activity, transferring phosphorus-containing groups

10 -id: GO:0009792
 name: embryonic development (sensu Metazoa)
 namespace: biological_process
 def: "The process whose specific outcome is the progression of an embryo over time, from zygote formation until the end of the embryonic life stage. The end of the embryonic life stage is organism specific and may be somewhat arbitrary; for mammals it is usually considered to be birth; for insects the hatching of the first instar larva from the eggshell. As in, but not restricted to, the multicellular animals (Metazoa, ncbi_taxonomy_id:33208)." [GOC:go_curators]
 is_a: GO:0009790 ! embryonic development

11 -id: GO:0005739
 name: mitochondrion
 namespace: cellular_component
 def: "A semiautonomous, self replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration." [ISBN:0198506732]
 is_a: GO:0043231 ! intracellular membrane-bound organelle
 is_a: GO:0044444 ! cytoplasmic part

12 -id: GO:0016021

name: integral to membrane

namespace: cellular_component

def: "Penetrating at least one phospholipid bilayer of a membrane. May also refer to the state of being buried in the bilayer with no exposure outside the bilayer. When used to describe a protein, indicates that all or part of the peptide sequence is embedded in the membrane." [GOC:go_curators]

is_a: GO:0031224 ! intrinsic to membrane

13 -id: GO:0004713

name: protein-tyrosine kinase activity

namespace: molecular_function

def: "Catalysis of the reaction: ATP + a protein tyrosine = ADP + protein tyrosine phosphate." [EC:2.7.1.112]

subset: gosubset_prok

is_a: GO:0004672 ! protein kinase activity

14 -id: GO:0004685

name: calcium- and calmodulin-dependent protein kinase activity

namespace: molecular_function

def: "Catalysis of the reaction: ATP + protein = ADP + O-phosphoprotein." [EC:2.7.1.123]

is_a: GO:0004683 ! calmodulin regulated protein kinase activity

15 -id: GO:0000074

name: regulation of progression through cell cycle

namespace: biological_process

def: "Any process that modulates the rate or extent of progression through the cell cycle." [GOC:go_curators]

is_a: GO:0051726 ! regulation of cell cycle

16 -id: GO:0005615

name: extracellular space

namespace: cellular_component

def: "That part of a multicellular organism outside the cells proper, usually taken to be outside the plasma membranes, and occupied by fluid." [ISBN:0198547684]

is_a: GO:0044421 ! extracellular region part

17 -id: GO:0003777

name: microtubule motor activity

namespace: molecular_function

def: "Catalysis of movement along a microtubule, coupled to the hydrolysis of a nucleoside triphosphate (usually ATP)." [GOC:mah, ISBN:0815316194]

comment: Consider also annotating to the molecular function term 'microtubule binding ; GO:0008017'.

is_a: GO:0003774 ! motor activity

18 -id: GO:0005215

name: transporter activity

namespace: molecular_function
 def: "Enables the directed movement of substances (such as macromolecules, small molecules, ions) into, out of, within or between cells." [GOC:ai, GOC:df]
 is_a: GO:0003674 ! molecular_function

19 -id: GO:0040010
 name: positive regulation of growth rate
 namespace: biological_process
 def: "Any process that increases the rate of growth of all or part of an organism." [GOC:mah]
 comment: Note that this term and its definition depart from the usual conventions for GO 'regulation' process terms; regulation of rate is not usually distinguished from regulation of extent or frequency, but it makes sense to do so for growth regulation.
 is_a: GO:0040009 ! regulation of growth rate
 is_a: GO:0045927 ! positive regulation of growth

20 - id: GO:0003989
 name: acetyl-CoA carboxylase activity
 namespace: molecular_function
 def: "Catalysis of the reaction: ATP + acetyl-CoA + HCO₃⁻ = ADP + phosphate + malonyl-CoA." [EC:6.4.1.2]
 is_a: GO:0016421 ! CoA carboxylase activity

21 -id: GO:0005524
 name: ATP binding
 namespace: molecular_function
 def: "Interacting selectively with ATP, adenosine 5'-triphosphate, a universally important coenzyme and enzyme regulator." [ISBN:0198506732]
 subset: gosubset_prok
 is_a: GO:0030554 ! adenyl nucleotide binding

22 -id: GO:0005829
 name: cytosol
 namespace: cellular_component
 def: "That part of the cytoplasm that does not contain membranous or particulate subcellular components." [ISBN:0198547684]
 is_a: GO:0044444 ! cytoplasmic part

23 -id: GO:0000786
 name: nucleosome
 namespace: cellular_component
 def: "A complex comprised of DNA wound around a multisubunit core and associated proteins, which forms the primary packing unit of DNA into higher order structures." [GOC:elh]
 is_a: GO:0043234 ! protein complex
 is_a: GO:0044427 ! chromosomal part
 relationship: part_of GO:0000785 ! chromatin

24 -id: GO:0005887
 name: integral to plasma membrane

namespace: cellular_component
 def: "Penetrating at least one phospholipid bilayer of a plasma membrane. May also refer to the state of being buried in the bilayer with no exposure outside the bilayer."
 [GOC:go_curators]
 is_a: GO:0016021 ! integral to membrane
 is_a: GO:0031226 ! intrinsic to plasma membrane

25 -id: GO:0005200
 name: structural constituent of cytoskeleton
 namespace: molecular_function
 def: "The action of a molecule that contributes to the structural integrity of a cytoskeletal structure." [GOC:mah]
 is_a: GO:0005198 ! structural molecule activity

26 -id: GO:0016020
 name: membrane
 namespace: cellular_component
 def: "Double layer of lipid molecules that encloses all cells, and, in eukaryotes, many organelles; may be a single or double lipid bilayer; also includes associated proteins." [GOC:mah, ISBN:0815316194]
 is_a: GO:0044464 ! cell part

28 -id: GO:0005743
 name: mitochondrial inner membrane
 namespace: cellular_component
 def: "The inner, i.e. lumen-facing, lipid bilayer of the mitochondrial envelope. It is highly folded to form cristae." [GOC:ai]
 is_a: GO:0019866 ! organelle inner membrane
 is_a: GO:0031966 ! mitochondrial membrane

29 – id: GO:0008151 cell growth and/or maintenance foi retirado (consulta em 10 de julho) passa a ser 'narrow synonym' de GO 50875 – Cellular physiological process

30 -id: GO:0004697
 name: protein kinase C activity
 namespace: molecular_function
 exact_synonym: "PKC activity" []
 xref_analog: EC:2.7.1.-
 is_a: GO:0001565 ! phorbol ester receptor activity
 is_a: GO:0004674 ! protein serine/threonine kinase activity

31 -id: GO:0003779
 name: actin binding
 namespace: molecular_function
 def: "Interacting selectively with monomeric or multimeric forms of actin, including actin filaments." [GOC:clt]
 is_a: GO:0008092 ! cytoskeletal protein binding

32 -id: GO:0004691
 name: cAMP-dependent protein kinase activity

namespace: molecular_function

alt_id: GO:0008602

def: "Catalysis of the reaction: ATP + a protein = ADP + a phosphoprotein; dependent on the presence of cAMP." [EC:2.7.1.37]

is_a: GO:0004690 ! cyclic nucleotide-dependent protein kinase activity

33 -id: GO:0004871

name: signal transducer activity

namespace: molecular_function

def: "Mediates the transfer of a signal from the outside to the inside of a cell by means other than the introduction of the signal molecule itself into the cell." [GOC:j], [ISBN:0198506732]

is_a: GO:0003674 ! molecular_function

34 -id: GO:0005622

name: intracellular

namespace: cellular_component

def: "The living contents of a cell; the matter contained within (but not including) the plasma membrane, usually taken to exclude large vacuoles and masses of secretory or ingested material. In eukaryotes it includes the nucleus and cytoplasm." [ISBN:0198506732]

is_a: GO:0044464 ! cell part

35 -id: GO:0001539

name: ciliary or flagellar motility

namespace: biological_process

def: "Locomotion due to movement of cilia or flagella." [GOC:hjd]

subset: gosubset_prok

is_a: GO:0006928 ! cell motility

37 -id: GO:0000226

name: microtubule cytoskeleton organization and biogenesis

namespace: biological_process

def: "The assembly and arrangement of cytoskeletal structures comprising microtubules and their associated proteins." [GOC:mah]

is_a: GO:0007017 ! microtubule-based process

38 -id: GO:0040032

name: post-embryonic body morphogenesis

namespace: biological_process

def: "The process by which the anatomical structures of the post-embryonic soma are generated and organized. Morphogenesis pertains to the creation of form." [GOC:ems, ISBN:0140512888]

comment: Note that this term was 'body morphogenesis (sensu Nematoda)'.

is_a: GO:0009886 ! post-embryonic morphogenesis

is_a: GO:0010171 ! body morphogenesis

39 -id: GO:0004692

name: cGMP-dependent protein kinase activity

namespace: molecular_function

def: "Catalysis of the reaction: ATP + a protein = ADP + a phosphoprotein; dependent on the presence of cGMP." [GOC:mah]
 is_a: GO:0004690 ! cyclic nucleotide-dependent protein kinase activity

40 -id: GO:0004842

name: ubiquitin-protein ligase activity
 namespace: molecular_function

def: "Catalysis of the reaction: ATP + ubiquitin + protein lysine = AMP + diphosphate + protein N-ubiquityllysine. Catalyzes the mediation of substrate recognition in ubiquitin-mediated protein degradation; binds directly to the substrate and its cognate E2 (ubiquitin conjugating enzyme)." [EC:6.3.2.19, PMID:9635407]

is_a: GO:0016881 ! acid-amino acid ligase activity

42 -id: GO:0003677

name: DNA binding
 namespace: molecular_function

def: "Interacting selectively with DNA (deoxyribonucleic acid)." [GOC:jl]

is_a: GO:0003676 ! nucleic acid binding

43 -id: GO:0004684

name: calmodulin-dependent protein kinase I activity
 namespace: molecular_function

def: "Catalysis of the reaction: ATP + protein = ADP + O-phosphoprotein." [EC:2.7.1.123]

is_a: GO:0004683 ! calmodulin regulated protein kinase activity

44 -id: GO:0005856

name: cytoskeleton
 namespace: cellular_component

def: "Any of the various filamentous elements within the cytoplasm of eukaryotic cells that remain after treatment of the cells with mild detergent to remove membrane constituents and soluble components of the cytoplasm. The term embraces intermediate filaments, microfilaments, microtubules, and the microtrabecular lattice. The various elements of the cytoskeleton not only serve in the maintenance of cellular shape but also have roles in other cellular functions, including cellular movement, cell division, endocytosis, and movement of organelles." [ISBN:0198547684]

is_a: GO:0043232 ! intracellular non-membrane-bound organelle

45 -id: GO:0007605

name: sensory perception of sound
 namespace: biological_process

def: "The series of events required for an organism to receive an auditory stimulus, convert it to a molecular signal, and recognize and characterize the signal. Sonic stimuli are detected in the form of vibrations and are processed to form a sound." [GOC:ai]

is_a: GO:0050954 ! sensory perception of mechanical stimulus

46 - id: GO:0000910

name: cytokinesis

namespace: biological_process

def: "The processes resulting in the division of the cytoplasm of a cell and its separation into two daughter cells. Cytokinesis usually occurs after growth, replication, and segregation of cellular components." [GOC:ai]

is_a: GO:0051301 ! cell division

47 -id: GO:0000165

name: MAPKKK cascade

namespace: biological_process

def: "Cascade of at least three protein kinase activities culminating in the phosphorylation and activation of a MAP kinase. MAPKKK cascades lie downstream of numerous signaling pathways." [GOC:mah, PMID:9561267]

is_a: GO:0007243 ! protein kinase cascade

48 -id: GO:0000788

name: nuclear nucleosome

namespace: cellular_component

def: "A complex comprised of DNA wound around a multisubunit core and associated proteins, which forms the primary packing unit of DNA in the nucleus into higher order structures." [GOC:elh]

is_a: GO:0000786 ! nucleosome

is_a: GO:0044454 ! nuclear chromosome part

relationship: part_of GO:0000790 ! nuclear chromatin

49 -id: GO:0004711

name: ribosomal protein S6 kinase activity

namespace: molecular_function

def: "Catalysis of the phosphorylation of ribosomal protein S6." [GOC:mah, PMID:9822608]

is_a: GO:0004674 ! protein serine/threonine kinase activity

50 -id: GO:0000082

name: G1/S transition of mitotic cell cycle

namespace: biological_process

def: "Progression from G1 phase to S phase of the mitotic cell cycle." [GOC:mah]

relationship: part_of GO:0051329 ! interphase of mitotic cell cycle

51 - id: GO:0006817

name: phosphate transport

namespace: biological_process

def: "The directed movement of phosphate into, out of, within or between cells." [GOC:krc]

is_a: GO:0015698 ! inorganic anion transport

52 -id: GO:0004736

name: pyruvate carboxylase activity

namespace: molecular_function

def: "Catalysis of the reaction: ATP + pyruvate + HCO₃⁻ = ADP + phosphate + oxaloacetate." [EC:6.4.1.1]

is_a: GO:0016885 ! ligase activity, forming carbon-carbon bonds

53 -id: GO:0004409

name: homoaconitate hydratase activity

namespace: molecular_function

def: "Catalysis of the reaction: 2-hydroxybutane-1,2,4-tricarboxylate = but-1-ene-1,2,4-tricarboxylate + H₂O." [EC:4.2.1.36]

is_a: GO:0016836 ! hydro-lyase activity

54 - id: GO:0003700

name: transcription factor activity

namespace: molecular_function

alt_id: GO:0000130

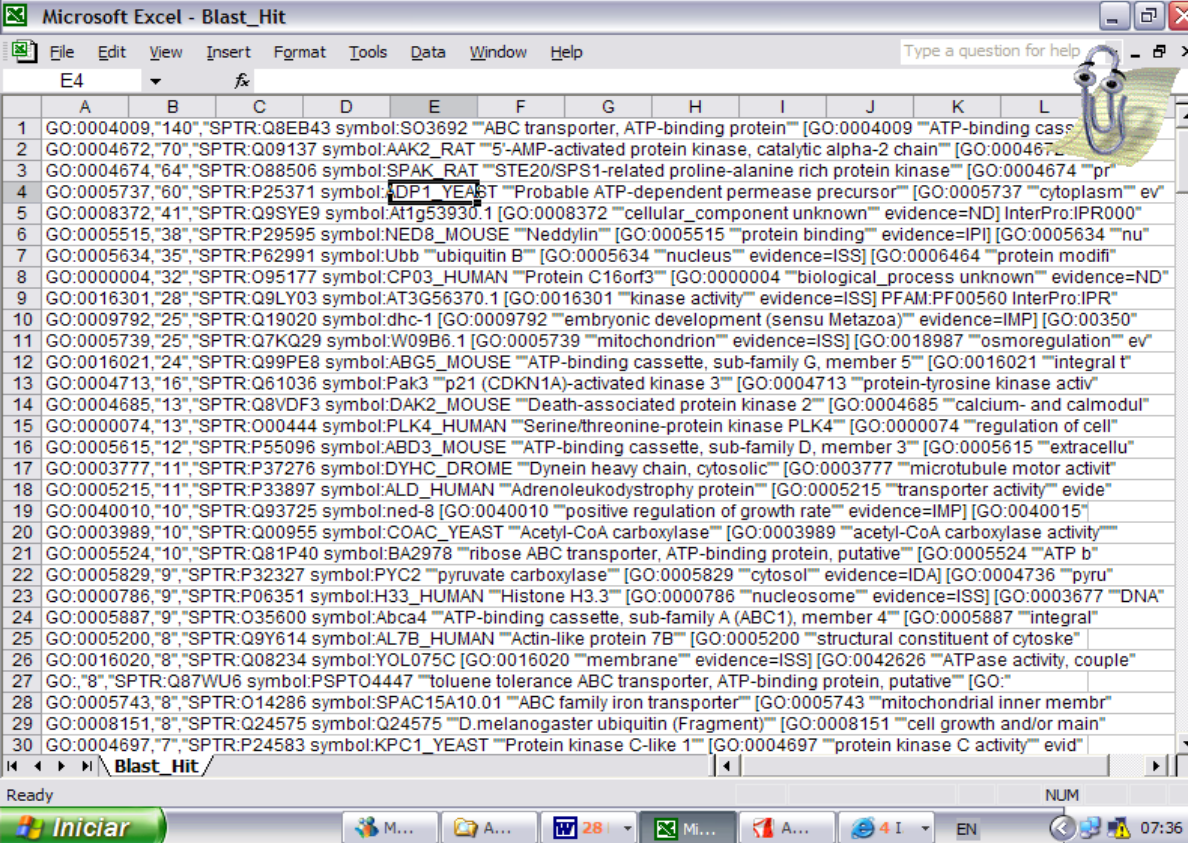
def: "The function of binding to a specific DNA sequence in order to modulate transcription. The transcription factor may or may not also interact selectively with a protein or macromolecular complex." [GOC:curators]

is_a: GO:0003677 ! DNA binding

is_a: GO:0030528 ! transcription regulator activity

ANEXO A

Lista de termos mais ocorrentes na base de dados genômicos dentro do projeto BioWebDB⁴⁰



	A	B	C	D	E	F	G	H	I	J	K	L
1	GO:0004009	"140"	"SPTR:Q8EB43	symbol:SO3692	"ABC transporter, ATP-binding protein"	[GO:0004009	"ATP-binding cass					
2	GO:0004672	"70"	"SPTR:Q09137	symbol:AAK2_RAT	"5-AMP-activated protein kinase, catalytic alpha-2 chain"	[GO:0004672						
3	GO:0004674	"64"	"SPTR:O88506	symbol:SPAK_RAT	"STE20/SPS1-related proline-alanine rich protein kinase"	[GO:0004674	"pr					
4	GO:0005737	"60"	"SPTR:P25371	symbol:ADP1_YEAST	"Probable ATP-dependent permease precursor"	[GO:0005737	"cytoplasm" ev					
5	GO:0008372	"41"	"SPTR:Q9SYE9	symbol:At1g53930.1	[GO:0008372	"cellular_component unknown"	evidence=ND	InterPro:IPR000				
6	GO:0005515	"38"	"SPTR:P29595	symbol:NED8_MOUSE	"Neddylin"	[GO:0005515	"protein binding"	evidence=IPJ	[GO:0005634	"nu		
7	GO:0005634	"35"	"SPTR:P62991	symbol:Ubb	"ubiquitin B"	[GO:0005634	"nucleus"	evidence=ISS	[GO:0006464	"protein modifi		
8	GO:0000004	"32"	"SPTR:O95177	symbol:CP03_HUMAN	"Protein C16orf3"	[GO:0000004	"biological_process unknown"	evidence=ND				
9	GO:0016301	"28"	"SPTR:Q9LY03	symbol:AT3G56370.1	[GO:0016301	"kinase activity"	evidence=ISS	PFAM:PF00560	InterPro:IPR			
10	GO:0009792	"25"	"SPTR:Q19020	symbol:dhc-1	[GO:0009792	"embryonic development (sensu Metazoa)"	evidence=IMP	[GO:00350				
11	GO:0005739	"25"	"SPTR:Q7KQ29	symbol:W09B6.1	[GO:0005739	"mitochondrion"	evidence=ISS	[GO:0018987	"osmoregulation" ev			
12	GO:0016021	"24"	"SPTR:Q99PE8	symbol:ABG5_MOUSE	"ATP-binding cassette, sub-family G, member 5"	[GO:0016021	"integral t					
13	GO:0004713	"16"	"SPTR:Q61036	symbol:Pak3	"p21 (CDKN1A)-activated kinase 3"	[GO:0004713	"protein-tyrosine kinase activ					
14	GO:0004685	"13"	"SPTR:Q8VDF3	symbol:DAK2_MOUSE	"Death-associated protein kinase 2"	[GO:0004685	"calcium- and calmodul"					
15	GO:0000074	"13"	"SPTR:O00444	symbol:PLK4_HUMAN	"Serine/threonine-protein kinase PLK4"	[GO:0000074	"regulation of cell"					
16	GO:0005615	"12"	"SPTR:P55096	symbol:ABD3_MOUSE	"ATP-binding cassette, sub-family D, member 3"	[GO:0005615	"extracellu"					
17	GO:0003777	"11"	"SPTR:P37276	symbol:DYHC_DROME	"Dynein heavy chain, cytosolic"	[GO:0003777	"microtubule motor activit					
18	GO:0005215	"11"	"SPTR:P33897	symbol:ALD_HUMAN	"Adrenoleukodystrophy protein"	[GO:0005215	"transporter activity" evide					
19	GO:0040010	"10"	"SPTR:Q93725	symbol:ned-8	[GO:0040010	"positive regulation of growth rate"	evidence=IMP	[GO:0040015				
20	GO:0003989	"10"	"SPTR:Q00955	symbol:COAC_YEAST	"Acetyl-CoA carboxylase"	[GO:0003989	"acetyl-CoA carboxylase activity"					
21	GO:0005524	"10"	"SPTR:Q81P40	symbol:BA2978	"ribose ABC transporter, ATP-binding protein, putative"	[GO:0005524	"ATP b"					
22	GO:0005829	"9"	"SPTR:P32327	symbol:PYC2	"pyruvate carboxylase"	[GO:0005829	"cytosol"	evidence=IDA	[GO:0004736	"pyru"		
23	GO:0000786	"9"	"SPTR:P06351	symbol:H33_HUMAN	"Histone H3.3"	[GO:0000786	"nucleosome"	evidence=ISS	[GO:0003677	"DNA"		
24	GO:0005887	"9"	"SPTR:O35600	symbol:Abca4	"ATP-binding cassette, sub-family A (ABC1), member 4"	[GO:0005887	"integral"					
25	GO:0005200	"8"	"SPTR:Q9Y614	symbol:AL7B_HUMAN	"Actin-like protein 7B"	[GO:0005200	"structural constituent of cytoske"					
26	GO:0016020	"8"	"SPTR:Q08234	symbol:YOL075C	[GO:0016020	"membrane"	evidence=ISS	[GO:0042626	"ATPase activity, couple"			
27	GO:	"8"	"SPTR:Q87WU6	symbol:PSPT04447	"toluene tolerance ABC transporter, ATP-binding protein, putative"	[GO:						
28	GO:0005743	"8"	"SPTR:O14286	symbol:SPAC15A10.01	"ABC family iron transporter"	[GO:0005743	"mitochondrial inner membr"					
29	GO:0008151	"8"	"SPTR:Q24575	symbol:Q24575	"D.melanogaster ubiquitin (Fragment)"	[GO:0008151	"cell growth and/or main"					
30	GO:0004697	"7"	"SPTR:P24583	symbol:KPC1_YEAST	"Protein kinase C-like 1"	[GO:0004697	"protein kinase C activity" evid					

⁴⁰ Resultado de uma busca nas bases de dados genômicos para identificar os termos mais frequentes nas bases, fornecida por Pablo Mendes no primeiro semestre de 2005.