

GT 2 : Organização do Conhecimento e Representação da Informação

ONTOLOGIAS DE DOMÍNIO: UM ESTUDO DAS RELAÇÕES CONCEITUAIS E SUA APLICAÇÃO

Luana Farias Sales

Instituto de Engenharia Nuclear

Maria Luiza de Almeida Campos

Universidade Federal Fluminense, Departamento de Ciência da Informação

Hagar Espanha Gomes

Consultora Independente

RESUMO

No âmbito da elaboração de Ontologias uma problemática que tem se colocado é a ausência de um padrão teórico-metodológico para sua elaboração. Propõe-se então uma sistematização das relações apresentadas nas literaturas da Ciência da Informação, da Terminologia, da Ciência da Computação e da Bioinformática. 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. 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. The study proposes systematization of relations found in the the literature from Information Science, Terminology, Computer Science and Bioinformatics. The systematization proposed aims at linking couple of categories (categorical relations) with relations properly (formal relations). 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

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.

Este projeto foi criado por Tim Berners-Lee enquanto trabalhava no laboratório europeu de Física CERN (Organisation Européenne pour la Recherche Nucléaire), 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. (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 a necessidade 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 a criação de um consórcio Internacional chamado W3C - World Wide Web Consortium.

A este consórcio se unem empresas, instituições acadêmicas, profissionais e cientistas com o objetivo de criar uma nova Web, a Web Semântica, cujo objetivo é 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

Moura (2002, p.2), no entanto, afirma que a interoperabilidade dos dados deve acontecer em três 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 “específica 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 semântico, eles retomam, uma proposta da Inteligência Artificial, que é a criação dos vocabulários controlados, com linguagem própria de máquina, denominados **Ontologias**.

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 presente pesquisa foi “como a Ciência da Informação poderia contribuir para elaboração de Ontologias?”.

A falta de consistência metodológica para a elaboração de ontologias foi apontada como problema por diversos autores (FERNANDEZ-LOPEZ, 1999, JONES et al, 1996). Entre os problemas ocasionados pela falta de consistência, 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, já que as relações entre conceitos, estabelecidas de forma coerente, vão permitir a criação de regras de inferência possibilitando, assim, uma interpretação formal mais precisa. Para isso, buscaram-se bases teóricas nas áreas de Ciência da Informação, Terminologia, Ciência da Computação e Bioinformática;

2 ONTOLOGIA: DEFINIÇÃO, COMPOSIÇÃO, APLICAÇÃO E TIPOS

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. 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, buscando, com isto, uma definição do objeto (MOREIRA, ALVARENGA E OLIVEIRA, 2004); TRISTÃO, FACHIN, ALARCON, 2004), porém, os autores sempre partem de definições dadas pela Ciência da Computação.

Algumas definições apresentadas na literatura de Ciência da Computação e principalmente na área de Inteligência Artificial, relacionam as *ontologias* a simples estruturas de conceitos. Chandrasekaran e Josephson (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 (1996, 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 apenas 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), porém, 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 (2002, p.123), que 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.

Sobre a definição do objeto “Ontologia”, este estudo destaca os seguintes pontos em relação às definições apresentadas: 1) Ontologias possuem conceituações que devem ser compartilhadas. 2) Ontologias devem ser descritas através de axiomas lógicos. 3) Ontologias devem ser processáveis por máquina.

Estes pontos esclarecem que: as ontologias devem possuir conceituações compartilhadas, ou seja, elas devem possuir um vocabulário estruturado com termos, definições e relacionamentos, que devem expressar um acordo comum entre seus usuários; elas devem ser descritas através de axiomas lógicos e suas definições devem possuir uma forma padrão, baseada em axiomas, que possam determinar a verdade das sentenças; e, ainda, ser processáveis por máquinas, ou seja, 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: **Conceitos** - que são idéias básicas sobre o que se pretende formalizar. Os conceitos podem ser classes de objetos, métodos, planos, estratégias, processos etc. **Classes e Subclasses** - que podem estar organizadas em uma taxonomia. **Relações** - que devem representar os tipos de interação entre as classes de um domínio. Estas relações são formalmente definidas como qualquer subconjunto dos produtos de um conjunto e são sempre binárias, como por exemplo: subclass_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. **Axiomas** - são teoremas que se declaram sobre as relações que devem cumprir todos 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; GRUNINGER, 1996, p.93)

Quanto às possibilidades de aplicação de ontologias, podem-se citar as seguintes categorias: Comunicação (USCHOLD, GRUNINGER, 1996), Especificação de Sistemas (USCHOLD; GRUNINGER, 1996; VICKERY, 1997; JASPER; USCHOLD, 1999; MOREIRA, 2002), Reutilização (USCHOLD; GRUNINGER, 1996; SWARTOUT et al., 1997; JASPER; USCHOLD, 1999), Engenharia de sistemas (USCHOLD; GRUNINGER, 1996), Confiabilidade (USCHOLD; GRUNINGER, 1996), Processamento de textos em Linguagem Natural (VICKERY, 1997; ALMEIDA; BAX, 2003), Classificação para ação (SOERGEL, 1999), Usabilidade (MOREIRA, 2002), Interoperabilidade (USCHOLD; GRUNINGER, 1996; VICKERY, 1997), Representação e Recuperação de Informação (SOWA, 2000; VICKERY, 1997).

Inúmeras são as tipologias registradas na literatura, algumas delas conflitantes. 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,

linguísticas/terminológicas, ontologias semi-informais, ontologias formais e ontologias rigorosamente formais.

Esta pesquisa utiliza um tipo específico de ontologia para análise de relações, que é a Ontologia de Domínio e pode ser definida como 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. A Ontologia de Domínio utilizada por este trabalho como campo empírico para análise das relações é a Gene Ontology (GO).

3 A GENE ONTOLOGY

A Gene Ontology (GO) faz parte de um grupo de Ontologias chamado de OBO (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.”

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 para quais processos biológicos estas funções contribuem?” 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.

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”.

Smith et al (2003, 2005) destacam um problema, que vem justificar a importância desta pesquisa, ou seja, a inconsistência nas relações entre conceitos no âmbito desta ontologia. Segundo eles “tais erros derivam-se de uma falta de atenção a princípios ontológicos.”

4 RELAÇÕES CONCEITUAIS NO ÂMBITO DA CIÊNCIA DA INFORMAÇÃO, DA TERMINOLOGIA E DA CIÊNCIA DA COMPUTAÇÃO E BIOINFORMÁTICA: A SISTEMATIZAÇÃO

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), na estrutura terminológica (USCHOLD, KING, 1995, FERNANDEZ et al, 1997), e na elaboração da taxonomia que deve compor a estrutura da terminologia (GUARINO 1995; SURE, STAAB, STUDER, 2002).

Sheth, Arpinar, Kashyap (2002) afirmam que os relacionamentos são fundamentais para a 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 na visão da Web Semântica.

A pesquisa que gerou este estudo apresentou uma revisão de literatura abrangendo as áreas de Ciência da Informação, Terminologia, Ciência da Computação e Bioinformática. A análise destas áreas mostrou, de uma maneira geral, que 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 de acordo com 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. 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 e 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 no desenvolvimento de linguagens documentárias. Neste sentido, estudos mais consistentes sobre relações conceituas são verificados, mais precisamente, no âmbito da elaboração de linguagens documentárias pós-coordenada, onde pode ser averiguada a existência de relações genéricas, partitivas e associativas

Na Terminologia a ênfase é dada ao desenvolvimento de terminologias e glossários, 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, é útil para expor em que sentido foi dado o 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, incluem-se as partitivas e as associativas.

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 venham sendo explicitadas através de verbos, verbos preposicionados, conjunções e outros. Porém, 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, é de grande benefício, pois liga classes de conceitos.

Já na parte formal não basta ligar as classes, mas é 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.

A área Biomédica, por sua vez, 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 já vêm sendo realizadas no sentido de promover o estabelecimento de relações entre conceitos de forma consistente.

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. Das 290 relações foi feita

uma análise do que os autores entendiam por cada uma, eliminado 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 Categoriais (total de 67) e Relações Formais (total de 59).

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

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 manifesta sempre da mesma forma. Isso faz com que não haja necessidade de sistematização das mesmas. Como tipo de relações genéricas encontramos na literatura os seguintes casos:

Quadro 1 – Sistematização das relações genéricas

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

As **relações partitivas**, porém, 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).

A sistematização das relações partitivas foi realizada da seguinte forma: primeiramente, foram selecionadas na listagem de relações (Apêndice A) 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:

Quadro 2 – Sistematização das relações partitivas

Relações Categoriais Partitivas	Relações Funcionais Partitivas
<ul style="list-style-type: none"> • Objeto-Constituinte • Constituinte-objeto 	<ul style="list-style-type: none"> Has_constituent Has_holo_constituent
<ul style="list-style-type: none"> • Objeto-unidade • Unidade-Objeto 	<ul style="list-style-type: none"> Has_unity Has_holo_unity
<ul style="list-style-type: none"> • Coleção-elemento • Elemento-coleção 	<ul style="list-style-type: none"> Has_element Has_collection
<ul style="list-style-type: none"> • Massa-porção • Porção-massa 	<ul style="list-style-type: none"> Has_portion Has_mass
<ul style="list-style-type: none"> • Objeto-material • Material-objeto 	<ul style="list-style-type: none"> Has_material Has_holo_material
<ul style="list-style-type: none"> • Área-lugar • Lugar-área 	<ul style="list-style-type: none"> Has_locate Has_holo_locate
<ul style="list-style-type: none"> • Zona-região 	<ul style="list-style-type: none"> Has_locate Has_holo_locate

• Região-zona	
• Caráter-atividade	Has_stage
• Atividade-caráter	Has_holo_stage
• Processo contínuo-ato	Has_stage
• Ato-processo contínuo	Has_holo_stage
• Processo descontínuo-fase	Has_stage
• Fase-processo descontínuo	Has_holo_stage
• Grupo-membro	Has_member
• Membro-grupo	Has_holo_stage
• Classe-Membro	Has_member
• Membro-classe	Has_holo_stage
• Substância-particular	Has_substance
• Particular-substância	Has_holo_substance
• Conjunto-subconjunto	Has_subset
• Subconjunto-Conjunto	Has_set
• Conjunto-elemento	Has_element
• Elemento-conjunto	Has_set
• Todo/sub-parte temporal	Has_subpart
• Sub_parte temporal - todo	Has_holo
• Ciência-objeto de estudo	Has_branch
• Objeto de estudo - ciência	Has_hole_branch

As **relações funcionais** 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.

A sistematização das relações categoriais se deu através do arranjo e combinação das categorias ranganathianas PMEST (personalidade, matéria, energia/processo, 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.

Essa distribuição serviu para categorizar as relações apresentadas na literatura, porém, como era esperado, as relações identificadas na literatura não esgotam as possibilidade 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 A). Para cada dupla de ocorrência foram selecionadas todas as possibilidades de relações formais, chegando assim, à seguinte sistematização:

Quadro 3 – Sistematização das Relações Funcionais

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 • Interacts with • Manifestation of • Analyses

	<ul style="list-style-type: none"> • Measures • Creates • Placed in • Located in • Adjacent to • 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	<ul style="list-style-type: none"> • Is a measure of
Medida – Processo	<ul style="list-style-type: none"> • 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"> • Property of • Counteragent of
Propriedade-processo	<ul style="list-style-type: none"> • Property of

Esta 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, devem se criar regras para quando usar cada relação.

Uma amostra da Gene Ontology, para análise das relações formais, resultou em 14 novas relações. Esse fato ocorreu porque, como já mencionado, as relações entre conceitos podem se manifestar 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 distintas, fazendo com que na elaboração dessas terminologias haja sempre a necessidade de se pensar quais relações são importantes para tal contexto.

5 CONSIDERAÇÕES FINAIS: APLICAÇÃO DO MODELO DE RELAÇÕES NO ÂMBITO DA GENE ONTOLOGY

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 nesta ontologia.

Outra 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 revelava apenas as categorias, já na Ciência da Computação e na Bioinformática pôde-se verificar a existência de outros modelo de relações, ou seja, relações formais, que revelavam, apenas, o tipo de relação existente entre essas duas categorias. Isto pode ser explicado pelo propósito em que cada relação foi criada, ou seja, para tesouros ou ontologias

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.

6 REFERÊNCIAS

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.

BERNERS-LEE, T; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, [S.l.], 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.

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

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.; 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.

DING, Y; 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, 2002.

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.

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

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

GUARINO, N. 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.

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.

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.

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.

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.

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.

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

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

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

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.

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

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

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.

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.

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

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.

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

W3C. **Semantic web**. 2005. Disponível em : <<http://www.w3c.org/2001/ws>>. Acesso em: 10 jul. 2006.

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