**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE COMUNICAÇÕES E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
INFORMAÇÃO**

JOLIZA CHAGAS FERNANDES

**O UNIVERSO E AS RELAÇÕES DE SIGNIFICAÇÃO DA
WEB: SEMIOSE NAS ONTOLOGIAS**

**SÃO PAULO
2012**

**O UNIVERSO E AS RELAÇÕES DE SIGNIFICAÇÃO DA
WEB: SEMIOSE NAS ONTOLOGIAS**

**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE COMUNICAÇÕES E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
INFORMAÇÃO**

JOLIZA CHAGAS FERNANDES

**O UNIVERSO E AS RELAÇÕES DE SIGNIFICAÇÃO DA
WEB: SEMIOSE NAS ONTOLOGIAS**

**Tese apresentada à Escola de
Comunicações e Artes, da Universidade de
São Paulo, como exigência para obtenção
do título de Doutor em Ciência da
Informação, Programa de Pós-graduação
em Ciência da Informação-Área de
Concentração Cultura e Informação.**

ORIENTADORA: Prof. Dra. NAIR YUMIKO KOBASHI

**SÃO PAULO
2012**

JOLIZA CHAGAS FERNANDES

**O UNIVERSO E AS RELAÇÕES DE SIGNIFICAÇÃO DA
WEB: SEMIOSE NAS ONTOLOGIAS**

Tese apresentada ao Programa de Pós-graduação da Escola de Comunicação e Artes da Universidade de São Paulo, na linha de pesquisa “Acesso à Informação”, em cumprimento das exigências para obtenção do título de Doutor.

Aprovada em 23 / 08 / 12

BANCA EXAMINADORA:

Orientadora: Profa. Dra. Nair Yumiko Kobashi

Examinador: Prof^o Dr. Marcelo dos Santos

Examinador: Prof. Dr. Marivalde Moacir Francelin

Examinadora: Prof^a Dr^a Maria Luiza Almeida Campos

Examinador: Prof^o Dr^o Carlos Xavier de Azevedo Netto

*Ao meu filho,
Benício Fernandes de Barros,
e ao meu marido,
Ronaldo Barros,
pelo que representam
em minha vida,
amor e esperança.*

AGRADECIMENTOS

A Deus, por sua infinita bondade e pelo dom da vida.

Ao meu marido, Ronaldo F. de Barros, pelo incansável apoio, pela paciência nas horas difíceis, pelo carinho nas horas certas, pelo companheirismo em todos os momentos..

A minha mãe, Luiza C. de Oliveira, pelo apoio incondicional em todas as horas, pela credibilidade que sempre depositou nos meus estudos, pela inspiração de luta pela vida que nos ensinou a trilhar os caminhos de maneira segura e digna.

Aos meus irmãos: Socorro, Zé Roberto, Fernando e Jadson que de uma maneira ou de outra me impulsionaram a continuar a caminhada.

A família Fernandes, na pessoa da Tia Maria, por terem proporcionado uma vida em família durante a minha estadia em São Paulo.

A Profª. Drª. Nair Yumiko Kobashi, orientadora, pela paciência, incentivo, confiança e pelos novos conhecimentos ofertados, o meu eterno agradecimento.

Aos professores Ana Maria Cintra e Marcelo dos Santos, pelas observações emitidas ao projeto, que contribuíram grandemente para construção deste estudo.

Aos meus professores do PPGCI: Ana Maria Cintra, Daisy Noronha, Antonio Gonzales Moreira, Fatima Tálamo, Johana Smith, por proporcionarem uma nova visão de conteúdos, leituras, críticas...

Aos colegas do curso: Carlos, Monica, Cibebe, Renato, Marivalde, pela amizade.

À UFMT, pelo apoio na liberação para capacitação do doutorado, etapa importante em minha vida.

Agradeço a todos os que por ventura não tenha lembrado, mas que de alguma forma contribuíram para esta etapa da minha vida e me ensinaram que preciso ter persistência e integridade para vencer.

Resumo

As dificuldades de acesso à informação na Web são decorrentes, dentre outras, da ausência ou inadequação do tratamento das informações. Entre as diversas propostas de filtragem de conteúdos de informação disponíveis na Web tem tido destaque a ferramenta denominada Ontologia. As potencialidades das Ontologias para lidar com problemas semânticos, sobretudo quando há grandes volumes de informação, parece promissor. Dessa forma, a reflexão teórica e a observação empírica sobre as Ontologias são importantes e necessárias para compreender seus aspectos semânticos e formais (lógicos). Assim, com base na teoria peirceana, mais especificamente a teoria do interpretante, o estudo focalizou as questões que envolvem as relações semânticas e formais das Ontologias com o fim de identificar traços que permitam sistematizar parâmetros teóricos e metodológicos de sua construção e avaliação. Para tanto, adotou-se como estudo de caso a ontologia Radlex, instrumento ontológico da área da Saúde disponível gratuitamente na Web. Para a análise foram considerados 615 termos selecionados a partir de quatro classes das principais categorias do Radlex, a saber: “class Image Observation Characteristic”, “class Pathophysiologic Process Modifier”, “class Anatomical Cluster” e “class Object”. De acordo com as observações, concluiu-se que não basta a existência de aplicativos automatizados relevantes para o bom funcionamento desses instrumentos; são necessárias ações direcionadas para as questões de produção de sentido, que reflitam as necessidades informacionais dos pesquisadores que os utilizam. No caso do Radlex, apesar de trazer em sua estrutura conceitual todos os elementos necessários e pertinentes a uma representação da informação, é necessário enriquecê-la com relações semânticas e formais adicionais sobre a Radiologia.

Palavras-chave: Ontologia; Radlex; Representação da Informação, Organização e Tratamento da Informação; Recuperação de informação

Abstract

The difficulties to access information on the web are due, among other things, to the absence or inadequacy of the techniques used in their organization. Among the various proposals for content filtering available on the Web there is a prominent tool called Ontology. The potential of ontologies to deal with semantic problems, especially when there are large volumes of information, looks promising. On the other hand, theoretical and empirical observation on Ontologies are important and necessary to deeply understand aspects of semantic and formal (logical) relationships between terms they contain. Thus, based on Peirce's theory, more specifically the Theory of the interpretant, the study focused on issues involving the formal and semantic relationships of ontologies, intending to identify aspects to draw up parameters for its construction and evaluation. To this end, it was adopted as a case study the Radlex Ontology, complex health area ontological instrument available for free on the Web. For analysis 615 terms were selected from four classes of Radlex main categories, namely: "class Image Characteristic Observation", "class Modifier Pathophysiologic Process", "class Anatomical Cluster" and "class Object". According to the observations, we concluded that the existence of automated applications is not sufficient; it is necessary actions directed to the issues of meaning production, reflecting the information needs of researchers who use the ontologies. In the case of Radlex, despite bringing the necessary conceptual structure to represent information, the majority are hierarchical. Synonymy relationships are absent. It is necessary to enrich it with additional formal and semantic relationships non Radiology.

Keywords: Ontology; Radlex Ontologie; Information representation; Information Processing, Information retrieval.

LISTA DE FIGURAS

Figura 1:	A arquitetura da Web Semântica	p. 14
Figura 2:	Abordagem ontológica simples	p. 47
Figura 3:	Abordagem ontológica múltipla	p. 48
Figura 4:	Abordagem ontológica híbrida	p. 49
Figura 5:	Módulos da Ferramenta Protégé 2000	p. 88
Figura 6:	Interface da Ferramenta Protégé 2000	p. 90
Figura 7:	Fluxo de desenvolvimento de um projeto Protégé	p. 92
Figura 8:	Interface Radlex Ontology	p. 98
Figura 9:	Visualização do Radlex pela árvore hierárquica com o <i>plugin</i> OntoViz	p. 99
Figura 10:	Continuantes e Ocorrentes	p. 100
Figura 11:	Visualização do termo “Flow direction” na árvore hierárquica	p. 117

LISTA DE GRÁFICOS

Gráfico 1:	Porcentagem de relações formais, definições e sinônimos	p. 120
------------	---	--------

LISTA DE QUADROS

Quadro 1:	Vantagens e desvantagens da Ontologia diferentes baseadas abordagens de integração	p. 49
Quadro 2:	Detalhe da Categoria “Radlex Entity”	p. 101
Quadro 3:	Detalhamento do termo “Anatomical plane”	p. 112
Quadro 4:	Quadro 5: Detalhamento do termo “Anatomical compartment”	p. 117
Quadro 5:	Detalhamento do termo “Blinded interpretation”	p. 119
Quadro 6:	Detalhes do termo “Anatomical Cluster”	p. 127
Quadro 7:	Detalhes do termo “Lungs”	p. 129

SUMÁRIO DA PESQUISA

1 INTRODUÇÃO	p.11
1.1 Problematização do Objeto, as Hipóteses de trabalho, as justificativas, os objetivos e o percurso da pesquisa	p.12
2 ASPECTOS DA REPRESENTAÇÃO DE INFORMAÇÃO EM AMBIENTES HIPERMÍDIA	p.20
2.1 Representação: Abordagem Peirceana: teoria e conceitos	p.23
2.2 Pragmática Linguística e informação contextual	p.32
2.3 Abordagens pragmáticas na C.I.	p.33
3 SOBRE ONTOLOGIAS	p.39
3.1 Definição e histórico	p.39
3.2 Elementos de Composição	p.42
3.3 Tipos de Ontologias	p.44
3.4 Metodologias de construção	p.50
3.5 Ferramentas para construções de Ontologias	p.54
4 LINGUAGENS FORMAIS DE CONSTRUÇÃO DE ONTOLOGIAS	p.62
4.1 Lógica: área de abordagem da representação formal de enunciados	p.62
4.2 Formalismo: conceito, histórico e técnicas	p.67
4.3 Linguagens Formais para Construção de Ontologias	p.71
5 ANÁLISE SEMÂNTICA DA ONTOLOGIA RADLEX	p.81
5.1 Percurso metodológico	p.81
5.2 Universo e amostra	p.82
5.3 Critérios de análises das relações semânticas	p.83
5.4 Apresentação da Ferramenta PROTÉGÉ-2000	p.87
5.5 Apresentação do Radlex Ontology	p.92
5.5.1 <i>Conceito e Histórico</i>	p.92
5.5.2 <i>A Plataforma Protégé no Radlex Ontology</i>	p.96
5.5.3 <i>Estrutura do Radlex Ontology</i>	p.97
5.5.4 <i>Análise dos dados Radlex</i>	p.111
5.5.5 <i>Apresentação dos resultados da análise</i>	p.116
6. Considerações finais	p.124
REFERENCIAS	p.134
APENDICE	p.141
ANEXOS	p.153

1 INTRODUÇÃO

Os problemas de acesso e recuperação das informações que circulam na Web são objeto de pesquisas de diversas áreas do conhecimento, dentre as quais a Ciência da Informação. Os problemas de acesso decorrem, em geral, da quantidade de informação e da inadequação ou mesmo ausência de tratamento das mesmas. Em relação a isso, Bougnoux (1993, p.11) argumenta que

...muita informação pode matar a informação, suscita evasões imaginárias, recusa de saberes e se choca de qualquer maneira com o 'segredo informacional' de cada um (um organismo só utiliza uma ínfima parte dos sinais que perpassam pelo seu meio ambiente).

Na mesma linha de raciocínio, Pierre Lévy (1997, p.24) acentua que

dispomos de poucas ferramentas para filtrar informações pertinentes, na perspectiva de estabelecer comparações de acordo com os significados e as necessidades sempre subjetivas, para nos movimentarmos no fluxo informacional.

Dentre as diversas propostas de filtragem de conteúdos de informações da Web pode ser destacada a Ontologia, ferramenta para atender de forma mais pertinente e adequada as necessidades de informação. Com efeito, em

muitas situações, os usuários desejam ter acesso não apenas às referências de documentos, mas sim à própria informação, ao conteúdo inscrito em variados tipos de suportes, portanto, a documentos relacionados a um determinado tema.

Em trabalho anterior (FERNANDES, 2005), havíamos constatado que os mecanismos utilizados na WEB são ainda pouco eficientes para identificar ou dar acesso aos conteúdos propriamente ditos. As pesquisas sobre as Ontologias vêm sendo realizadas justamente para promover melhor compreensão desses problemas, como também para a busca de melhores formas de organizá-las e recuperá-las. A perspectiva de aprimoramento das buscas está, a nosso ver, estreitamente vinculada à forma de organização de conteúdos. Dito de outro modo, resultados consistentes de busca são dependentes, em larga medida, do modo como os conteúdos estão organizados.

A compreensão das Ontologias requer investimentos de diferentes naturezas. Certos aspectos podem ser explorados no interior da Ciência da Informação, especificamente sob a perspectiva da Organização da Informação e do Conhecimento (OIC), com o propósito de criar metodologias adequadas de elaboração e avaliação desses instrumentos. É este o foco desta pesquisa. Para isso, consideramos fundamental empreender discussões teóricas sobre a representação, a significação e a formalização de conteúdos de informação. São estas as questões que foram exploradas nesta pesquisa, tendo como objeto específico as relações semânticas presentes em Ontologias.

1.1 Problematização do objeto, as Hipóteses de trabalho, as justificativas, os objetivos e o percurso da pesquisa.

Tradicionalmente, a organização e o tratamento da informação e do conhecimento (OIC) são realizados por meio de instrumentos terminológicos que procuram controlar a significação, de modo a eliminar a ambiguidade

provocada pela polissemia. Na WEB, a informação é, via de regra, indexada com unidades da linguagem natural, processo que torna vagos os significados das mesmas. É necessário lembrar que os indivíduos representam informação de acordo com contextos culturais específicos, o que nem sempre coincide com a forma de estruturação adotada na *WEB*. Este fato vem sendo discutido por pesquisadores da área, que propõem o uso de descritores¹ aptos a mudar a estrutura organizacional da rede (BUTLER, 2000)².

Há atualmente uma grande variedade de programas, provedores e produtores de serviços e repositórios ou sites temáticos que facilitam a recuperação de conteúdos de áreas especializadas, na WEB. Entretanto, as dificuldades de recuperação persistem em razão, ao menos, da não compreensão dos processos de significação e de representação sîgnicas por parte dos produtores de repositórios e dispositivos de informação.

O processo de significação de um signo ocorre por meio do interpretante. As interpretações não são subjetivas; ao contrário, refletem uma propriedade objetiva que o signo possui em si mesmo e que se projeta nos indivíduos (PEIRCE, 1977). É uma propriedade em constante construção, “que varia nos contextos de interlocução e dentro de instâncias culturais distintas, propiciando a elaboração de interpretações sobre o que está sendo representado” (AZEVEDO NETTO, 2002, p.57). Assim, o interpretante constitui um significado primeiro do signo; portanto, não há outro modo de se entender um signo a não ser pelo seu interpretante.

Nesse contexto, Pickler (2007) sinaliza a proliferação de projetos para dotar as máquinas de ferramentas ditas inteligentes, capazes de raciocinar e inferir os significados dos conteúdos dos documentos armazenados em bancos de dados. Uma dessas tecnologias são as Ontologias. No contexto da Web, não se utiliza o termo Ontologia no sentido filosófico (estudo do ser e da existência). Ela é uma ferramenta para atribuir sentido e significado a

¹ Os descritores, são unidades de linguagens documentárias, se distinguem das palavras da língua (unidades do léxico que representam apenas propriedades, nunca entidades, se constituindo em predicados livres sem um valor específico de referência), Os descritores, conforme Le Guern (1984), são unidades portadoras de informação.

² Divulgada em documento eletrônico na WEB

determinados termos, em dados contextos, de modo a precisar os conteúdos dos documentos armazenados (PICKLER, 2007). Para Davies (2003), uma Ontologia é um instrumento que procura estabelecer um vocabulário de representação para descrever uma realidade, capturando conceitos e suas respectivas relações num domínio específico; além disso, estabelece certas regras, os axiomas, para restringir a interpretação dos conceitos. Trata-se, portanto, de uma forma de lidar com a representação de recursos informacionais por meio de relações entre conceitos.

Ribeiro (2008), referindo-se à estrutura da Web semântica proposta pela WC3, explica que as ontologias surgiram justamente para compor uma das camadas centrais da arquitetura da Web Semântica, conforme figura abaixo.

Figura 1: A arquitetura da Web Semântica

Fonte: Ribeiro. Arquitetura da Web Semantica, 2012

- A camada **Unicode / URI** permite a interoperabilidade em relação à codificação de caracteres e ao endereçamento e nomeação de recursos da Web Semântica.

- A camada **XML / Namespace / XML Schema** permite a interoperabilidade em relação à sintaxe de descrição de recursos da Web Semântica.

- A camada **RDF / RDF Schema** fornece um *framework* para representar informação (metadados) sobre recursos.

- A camada **Ontologia** fornece suporte para o desenvolvimento de vocabulários e para processar e integrar a informação existente, de modo a evitar problemas de indefinição ou conflito de terminologias.

- A camada **Lógica** fornece suporte para a descrição de regras para expressar as relações entre os conceitos de uma Ontologia.

- As camadas **Prova** e **Confiança** fornecem suportes para a execução das regras, além de avaliar a correção e a confiabilidade dessas operações.

Segundo Moura (2009), as Ontologias, com sua noção de organização de conteúdos para promover a interoperabilidade entre sistemas, proporcionam um ambiente adequado ao fluxo informacional e ao trabalho colaborativo em rede, o que vem exigindo transformações também nas lógicas de tratamento da informação.

Os estudos mais abrangentes sobre as Ontologias vêm sendo realizados na área da Ciência da computação, cujas propostas voltam-se, em sua maioria, para a criação de estruturas de Ontologias específicas de áreas do conhecimento e sua formalização. Na Ciência da informação, a maioria dos estudos está voltada para o uso das Ontologias em ambientes virtuais de informação, sendo pouco frequentes os estudos sobre as linguagens de construção dessas ontologias e sobre os tipos de relações de significação nelas presentes. Ora, se as Ontologias visam ao entendimento do vocabulário

utilizado para representar um determinado conteúdo é essencial compreender esse sistema de representação e significação.

Com base nas reflexões acima, foram investigadas, nesta pesquisa, as relações de significação presentes em Ontologias, com o objetivo de contribuir para consolidar os conhecimentos necessários ao desenvolvimento da WEB semântica. O foco da pesquisa é a análise dos aspectos lógicos e semânticos das Ontologias, cujo objetivo é dar resposta aos seguintes questionamentos: como se dá o processo de representação na construção de significados nas Ontologias? Que elementos formais estão presentes na construção dessas representações? Quais ou qual é a teoria que poderá auxiliar o processo de busca de respostas para essas questões?

As potencialidades das Ontologias para lidar com problemas semânticos, sobretudo quando há grandes volumes de informação, parecem promissoras. A nosso ver, a reflexão teórica sobre este objeto deve ser acompanhado de observação empírica para identificar seus aspectos semânticos e formais (lógicos). É necessário, nesse sentido, problematizar as formas de inserir aspectos semânticos e lógicos na organização e recuperação da informação, processos inerentes à construção de Ontologias. Tendo em vista os limites das linguagens lógicas para expressar significados, pode-se inferir que os aspectos semânticos investidos nas Ontologias são também limitados. A teoria semiótica de Peirce mostra ser adequada para orientar as reflexões sobre este objeto porque ela considera a representação e a significação como processos lógicos e semânticos:

Em um sentido geral, a lógica é, como acredito ter mostrado, apenas um outro nome para semiótica, a quase-necessária, ou formal doutrina dos signos¹. Descrevendo a doutrina como quase-necessária” ou formal, quero dizer que observamos os caracteres de tais signos e a partir dessa observação, por um processo a que não objetarei denominar Abstração...[...] Quanto a esse processo de abstração, ele é, em si mesmo, uma espécie de observação (PEIRCE, 1977, P. 45)

Com base nas questões acima, apresentamos as hipóteses de trabalho desta pesquisa:

- a) os constituintes das Ontologias (termos) são semelhantes às presentes em vocabulários controlados do tipo tesauro.
- b) as relações entre os signos das Ontologias são representados em linguagem formal (lógica), distintas, portanto, das formas utilizadas em vocabulários controlados.
- c) os aspectos semânticos investidos nas Ontologias são limitados tendo em vista a abordagem lógica a que são submetidos seus constituintes (signos), tanto quanto as relações entre os signos.

Tendo em vista as hipóteses de trabalho acima, recorreremos à Teoria da Representação de Peirce para “observar” os signos e suas relações nas ontologias. De fato, os tipos de signos propostos por Peirce, as relações de significação, a semiose, suas reflexões sobre a representação e a Teoria dos interpretantes são categorias que permitem observar adequadamente diferentes sistemas de significação, no caso, as ontologias. Peirce se preocupa com os sistemas de significação, discute as questões terminológicas da atividade científica e adota visão pragmática ao se referir aos interpretantes. A dinâmica dos sistemas signicos são objeto de discussão específica, tanto quanto o processo de semiose, aspectos importantes das linguagens de representação de informação. A partir dessa tônica conceitual, Peirce afirma que o signo é um elemento triádico por natureza, comportando três dimensões, a princípio denominadas de qualidade, relação e representação e, mais tarde, de primeiridade, secundidade e terceiridade (SANTAELLA, 2005). Esse plano de discussão amplo dos sistemas de significação, relacionados ao uso dos signos na comunicação são mais potentes, a nosso ver, do que a abordagem estrutural da língua, de Saussure. Como se sabe, Saussure (2006), propõe a Linguística como o estudo do sistema linguístico, deixando de lado a fala e os falantes, ou seja, “tira da língua a fala e seus usuários”, estando portanto ausentes desta perspectiva as questões pragmáticas de uso da língua (SOUZA, 2010, p.114).

Com base nos problemas e hipóteses de trabalho acima apresentados, enunciamos os objetivos da pesquisa:

Objetivo Geral

- Sistematizar parâmetros teóricos e metodológicos de construção de Ontologias, com base na semiose peirceana.

Objetivos Específicos

- Refletir sobre as teorias da representação na construção de Ontologias, tomando por base a abordagem da semiose peirceana – em particular a noção de interpretante.
- Identificar e discutir as relações semânticas e formais presentes em Ontologias, tomando por base a teoria da semiose peirceana.

Apresentamos, a seguir, o percurso de investigação para a construção desta tese, expostos nos seguintes capítulos:

- Capítulo 1 – apresenta o problema e a problemática da pesquisa, as justificativas, as hipóteses de trabalho e os objetivos.
- Capítulo 2 – apresenta a fundamentação teórica sobre a representação da informação em ambientes hipermídia, incluindo na discussão as teorias e os conceitos relacionados ao tema, com enfoque especial na visão pragmática de Peirce.
- Capítulo 3 - apresenta o objeto de estudo desta tese, as ontologias. São explanadas as questões que envolvem sua definição, histórico, elementos de composição, tipos de ontologias, metodologias e ferramentas de construção.

- Capítulo 4 - são apresentadas as linguagens formais de construção de ontologias, baseadas em referenciais lógicos.
- Capítulo 5 – apresenta os aspectos metodológicos da pesquisa: as técnicas de pesquisa, o corpus e os critérios utilizados na sua definição.
- Capítulo 6 - é dedicado à análise do corpus, cujos desdobramentos incluem a apresentação da ferramenta Protegé 2000 e a ontologia de domínio Radlex.
- Capítulo 7 - apresenta a discussão dos resultados e as considerações finais: os resultados alcançados, as contribuições do estudo para a área de Organização da Informação e do Conhecimento, as limitações do estudo e as possíveis extensões do trabalho.

2 ASPECTOS DA REPRESENTAÇÃO DE INFORMAÇÃO EM AMBIENTES HIPERMÍDIA

Um dos elementos constitutivos do regime de informação (González de Gomez, 2003; Foucault, 1996; Frohmann, 1995), na sociedade atual, são a informação e as tecnologias da informação. Elas são recursos fundamentais para promover o aumento da produtividade, sendo, portanto, agentes importantes da mundialização da economia, do capital, da cultura e de poder.

Esse cenário é marcado por iniciativas para propor as formas mais eficientes de processar e disseminar conhecimentos, em particular os mecanismos de organizar, acessar, recuperar e interpretar conteúdos disponíveis na rede. Esse cenário motiva a área da organização da informação a se preocupar com a construção de modelos conceituais, bases de conhecimentos, modelagem de sistemas de informação e Ontologias (CAMPOS, 2001).

O conceito de representação é central nas discussões atuais da área da Ciência da informação. Em sentido amplo, porém, é utilizado em outros campos: Ciências cognitivas, Psicologia, Direito, na Filosofia, Sociologia e mais recentemente na Ciência da Computação. Nesses contextos, são utilizados, entre outros, os conceitos de representação social, representação processual, representação mental, representação do conhecimento e representação da informação.

A representação da informação ou do conhecimento é abordada segundo variadas perspectivas e linhas de raciocínio. Destaca-se, entre as abordagens, uma ciência que tem por objetivo estudar a ação dos signos linguísticos, a Semiótica. O termo tem origem grega com a denominação

semeiôtikos, cujo significado literal é a ótica dos sinais e o conceito a “ciência das leis necessárias gerais dos signos e, especialmente, dos símbolos” (PEIRCE, 1977, p.29)

De maneira mais abrangente, Santaella (1994, p.13) define a Semiótica como:

a ciência que tem por objeto de investigação todas as linguagens possíveis, ou seja, que tem por objetivo o exame dos modos de constituição de todo e qualquer fenômeno como fenômeno de produção de significação e de sentido. (...) Nos fenômenos, sejam eles quais forem – uma nesga de luz ou um teorema matemático, um lamento de dor ou uma idéia abstrata da ciência -, a Semiótica busca divisar e deslindar seu ser de linguagem, isto é, sua ação de signo. Tão só e apenas. E isso já é muito.

Tal ciência, nascida dos estudos filosóficos do teórico norte-americano Charles Sanders Peirce, em meados do sec. XX, permeia a “produção social de significado com base em sistemas signicos” (TRASK, 2008, p.263), isto é, estuda os fenômenos culturais e sociais para fundamentar os sistemas de significação.

No entanto, a Semiótica nem sempre foi o único termo utilizado para designar a teoria dos signos. A Semiologia, surgida na Europa, é outra denominação atribuída à teoria dos signos. Nöth esclarece essa disputa:

No nosso século, o termo semiologia ficou ligado à tradição semiótica fundada no quadro da linguística de Ferdinand de Saussure e continuada por semiotistas como Louis Hjelmslev ou Roland Barthes. Sob essas influências, a semiologia permaneceu durante muito tempo como o termo preferido nos países românicos, enquanto autores anglófonos e alemães preferiram o termo semiótica. Alguns semiotistas, porém, começaram a elaborar distinções conceituais entre semiologia e semiótica: semiótica, designando uma ciência mais geral dos signos, incluindo os signos animais e da natureza, enquanto semiologia passou a referir-se unicamente à teoria dos signos

humanos, culturais e, especificamente, textuais. (NÖTH, 2009, p.23)

De acordo com esse autor houve, em 1969, a tentativa de encerrar essa rivalidade com a criação da Associação Internacional da Semiótica, momento em que Roman Jakobson decidiu optar pelo termo Semiótica para designar o território das investigações sobre a semiologia e a semiótica geral.

No entanto, é importante ressaltar que, mesmo tendo havido a tentativa de fusão, ainda hoje coexistem essas duas linhas que regem a área da representação da informação: a linha linguística saussureana e a semiótica peirceana. Embora, ambas tragam em seu bojo os princípios linguísticos necessários para as atividades de representação da informação, comportam metodologias e enfoques distintos.

Na linha de Saussure (2006), o estudo dos signos se baseia em dois elementos principais: o significante, que diz respeito à parte física do signo, ou seja, à forma ou à sintaxe; e o significado, que se relaciona ao conteúdo. Para Saussure, o signo deve ser visto por ele mesmo: ele “tira da língua a fala e os seus usuários” (SOUZA, 2010, p.114).

Na linha peirceana, o plano de discussão é mais amplo, já que se considera o signo a partir de concepções de sociedade e de comunicação, colocando o usuário como principal agente da representação, esclarecendo que não há pensamento sem signo e vice-versa. De fato, “pensar é significar” (PEIRCE, 1977). A partir dessa tônica conceitual, Peirce conclui que o signo é um elemento triádico por natureza, comportando três dimensões, a princípio denominando-as de qualidade, relação e representação e mais tarde de primeiridade, secundidade e terceiridade. (SANTAELLA, 2005).

Nessas três dimensões podem-se encontrar as seguintes definições (SANTAELLA, 2011, p.9-11):

Primeiridade: é tudo o que está imediatamente presente à consciência de alguém, é tudo aquilo que está na sua mente no instante presente. (...) Nessa

medida, o primeiro é presente e imediato, ele é inicialmente, original, espontâneo e livre, ele precede toda síntese e toda diferenciação. Primeiridade é a compreensão superficial de um objeto.

Secundidade: a qualidade ou primeiridade é apenas uma parte do fenômeno, visto que, para existir, a qualidade tem que estar encarnada numa matéria. (...) Assim sendo, secundidade é quando o sujeito lê com compreensão e profundidade o seu conteúdo.

Terceiridade: corresponde à camada de inteligibilidade, ou pensamento em signos, através da qual representamos e interpretamos o mundo. (...) Ou seja, o indivíduo conecta à frase a sua experiência de vida, fornece à oração um contexto pessoal, providenciando assim uma interpretação final do signo recebido.

A partir dessas três categorias, Peirce construiu sua teoria da representação, colocando o homem, a matéria e os signos como elementos primordiais da produção de sentidos.

2.1 Representação: Abordagem Peirceana: teoria e conceitos

A construção de um conceito geral de representação se deu no momento da fundação do conceito de sociedade estabelecido pela teoria social moderna (MARTELETO, 2002), imbuído do pensamento sociológico e antropológico. No campo da Ciência da Informação, a representação é inerente ao processo de organização e circulação social da informação. Nesta área, os estudos apresentam abordagens teóricas e práticas variadas, ora focando aspectos cognitivos, ora linguísticos ou os aspectos comunicacionais, visitando domínios que vão desde o conhecimento em si, passando pelos sistemas de informação, indo até a relação dos usuários com os repertórios de informação (MARTELETO, 2002, p.03). O termo representação vem, pois, ganhando força no cenário científico com a explosão informacional, fenômeno que expandiu de forma crescente e irreversível todos os tipos de publicações, facilitando o armazenamento e a disseminação de dados no espaço virtual.

O contexto tecnológico informacional levantou uma preocupação: como tratar e organizar tamanha massa documental? Essa preocupação se tornou ainda maior com a popularização da internet, que ganhou status de produtora e difusora de informação para o cidadão comum, fazendo deste, ao mesmo tempo, um produtor e um disseminador de informação. Nesse período, estudos das áreas voltadas para a disseminação da informação, sobretudo computação e linguística, ganharam força. Teóricos como Luhn (1958) e Gardin (1969), entre outros, marcaram presença na área da organização do conhecimento dando contribuições significativas para as atividades de indexação e resumos, sobretudo Luhn que visava “modificar radicalmente os métodos tradicionais de armazenamento, tratamento e recuperação da informação” (apud KOBASHI, 1996).

Dentro da concepção peirceana, independentemente do domínio de atuação, “representar é estar em lugar de, isto é, estar numa tal relação com o outro que, para certos propósitos, é considerado por alguma mente como se fosse outro (PEIRCE, 1977, p.61)”. Santaella (2005, p.31) explica que, na visão peirceana, “qualquer coisa que substitui uma outra coisa para algum intérprete é uma representação ou signo”. Esse conceito pode designar o processo que coloca em correspondência pelo menos dois elementos, “de forma que um (o representante ou signo) repita, substitua ou apresente de outra forma o outro (o representado ou objeto). Designa também um aspecto do resultado desse processo, no caso, como o único elemento representante, qualquer que seja a sua natureza” (GRAND dictionnaire de la psychologie, 1991, p.596).

Peirce (1976) consolida sua primeira fase de estudos sobre o termo representação, buscando fundamentá-lo como a única forma de entender a sociedade, a cultura, o homem, e, conseqüentemente, todo conhecimento produzido a partir disso. Para ele, a “representação é fundamental ao pensamento, à ação, à percepção e emoção humanas”. (SAVAN, 1994 apud SANTAELLA, 2005, p.31).

Nessa primeira fase, a representação é sinônimo de signo, desconsiderando toda e qualquer concepção externa ao termo. Percebendo essa lacuna, Peirce desenvolveu o conceito, afirmando que a representação é

um processo de apresentação de um objeto por meio de um signo, estabelecendo relação direta entre o signo e o objeto, descartando totalmente a identidade anteriormente estabelecida. Nesse momento, Peirce introduz o *representamen* como um termo que delimita os conceitos de representação e signo: “quando é desejável distinguir entre aquilo que representa e o ato relação de representar, o primeiro pode ser chamado de *representamen* e o último de representação” (NÖTH apud SANTAELLA, 1997, p.65).

Assim, adota-se aqui o conceito de representação com viés peirceano, como sendo “estar para”, quer dizer, “algo que está numa relação tal com um outro que, para certos propósitos, ele é tratado por uma mente como se fosse aquele outro” (NÖTH, 1997, p.66). Esse conceito se firma na semiótica, aproximando-se do conceito linguístico de Benveniste. Para este autor, a linguagem está no lugar de outra coisa, ou seja, “a função do signo é a de representar, de tomar o lugar de outra coisa, evocando-a a título de seu substituto” (BENVENISTE apud KOBASHI, 1996, p.11).

Peirce define signo como “alguma coisa que representa algo para alguém [...] um *representamen* é o primeiro correlato de uma relação triádica [...] ou qualquer coisa que, de um lado, é assim determinado pelo objeto e, de outro, assim determina uma idéia na mente de uma pessoa” (PEIRCE apud SANTAELLA, 1997, p.23-27). Desse modo,

Um signo intenta representar, em partes, pelo menos, um objeto que é, portanto, num certo sentido, a causa ou o determinante do signo, mesmo que o signo represente o objeto falsamente. Mas dizer que ele representa seu objeto implica que ele afeta uma mente de modo que, de certa maneira, determina, naquela mente, algo que é mediatamente devido ao objeto. Essa determinação da qual a causa imediata ou determinante é o signo e da qual a causa mediata é o objeto pode ser chamada de interpretante. (PEIRCE, 1976 apud SANTAELLA, 2005, p.43)

O signo, neste contexto, é o sujeito principal do processo, sendo o objeto referenciado um elemento apresentado pelo signo, seja essa apresentação verdadeira ou não. Assim, o objeto pode ser uma coisa peculiar,

que já tenha existido ou que se espera venha a existir. Esta gama de variações que a noção pode abranger deixa clara a complexidade do conceito, muito embora o objeto seja um referente. É ele que determina a representação, não existindo sem ele o signo e sem signo não existe representação. O Interpretante é, pois, um terceiro elemento do conceito, que se funde aos dois primeiros, complementando-os e ligando-os de forma solidária.

Santaella (2004, p.10) argumenta que para a melhor compreensão dessas definições é importante considerar as três principais relações que as regem: a significação, a objetivação e a interpretação.

- Da relação com o signo consigo mesmo, isto é, da natureza do seu fundamento, ou daquilo que lhe dá capacidade de funcionar como tal, que, (...) pode ser sua qualidade, sua existência concreta ou seu caráter de lei, advém uma teoria das potencialidades e limites da **significação**.
- Da relação com o fundamento com o objeto, ou seja, com aquilo que determina o signo e que é, ao mesmo tempo, aquilo que o signo representa e ao qual se aplica, e que pode ser tomado em sentido genérico como o contexto do signo, extrai-se uma teoria da **objetivação**, que estuda todos os problemas relativos à denotação, à realidade e referência, ao documento e ficção.
- Da relação com o fundamento com o interpretante, deriva-se uma teoria da **interpretação**, com as implicações quanto aos seus efeitos sobre o intérprete, individual ou coletivo.

Esta autora salienta o ponto central da abordagem peirceana, a que relaciona a produção de sentidos dos signos ao pensamento humano, ou, aos aspectos cognitivos, afirmando que qualquer coisa que esteja presente na mente humana pode ser considerada como signo, pois tem potencialidade de significação. Nessa perspectiva ele amplia a noção de signo, considerando

emoção, reação, ação ou qualquer sentimento como elemento sgnico, bastando que sejam externalizados. Como tudo parte da cognio humana, tudo passvel de interpretao, o que confere qualidades e valores aos signos originados.

As interpretaes so inerentes ao terceiro elemento sgnico, o interpretante, que mantm relao estreita com o signo, ou representamen, e seu objeto. Na viso peirceana, esta relao mais estreita do que se apresenta, uma vez que existe a supremacia do objeto sobre o signo e do signo sobre o interpretante. Segundo Santaella, a supremacia se define da seguinte forma: o objeto algo diverso do signo, diversidade que o signo no pode apagar, o que determina a predominncia do objeto. Por isso, o signo no substitui o objeto, ao contrrio, determinado por ele, podendo apenas estar no lugar dele e indic-lo para a idia que ele mesmo produz. Esta relao de supremacia ocorre da mesma forma na determinao do interpretante pelo signo, embora com maior complexidade.

Considerando a importncia do pensamento humano na teoria de Peirce, pode-se entender a posio de destaque que ele atribui ao interpretante, sendo esta a parte mais extensa de sua abordagem.

O fato que tanto o objeto quanto o interpretante so partes constitutivas do signo. No entanto, este s pode ser definido se existir uma relao eficaz entre seu interpretante e o objeto, e justamente o interpretante que d as condies necessrias para a produo de sentidos em relao ao objeto, conseqentemente, para a definio eficaz do signo. Santaella (2002, p.23) esclarece esse fenmeno mencionando que “o interpretante o efeito interpretativo que o signo produz em uma mente real ou meramente em potencial.”

Na reflexo de Santaella, o interpretante se projeta neste processo no como uma interpretao subjetiva de cada indivduo, mas como uma propriedade objetiva que o signo possui em si mesmo e que se projeta em cada indivduo. Essa propriedade objetiva exterior ao signo, proveniente da

cultura e dos saberes acumulados nos grupos sociais, de geração para geração, que são eles mesmos signos.

Nesse sentido, Peirce coloca o interpretante numa relação de “terceridade”, explicando que “na sua forma genuína, a Terceiridade é uma relação triádica que existe entre um signo, seu objeto e um pensamento interpretador, ele mesmo um signo” (SANTAELLA, 1997, p.87). Portanto, observa-se que o signo interpretante constitui um significado primeiro e externo ao signo, e, sendo assim, não existe nenhum modo de se entender um signo a não ser pelo seu interpretante.

O interpretante é definido sob três aspectos: o imediato, o dinâmico e o final (SANTAELLA, 2002, CARDOSO, 2008). Essa classificação consiste no processo triádico indissolúvel de significação, apontado por Peirce (1977) como os três graus de clareza para compreender o signo.

Entretanto, no entendimento de Santaella (2002), a teoria só é explanável nas duas primeiras classificações, haja vista que para ela o interpretante final é inatingível. Mesmo assim, tentar-se-á uma explicação dos três níveis de interpretação porque o que interessa aqui é verificar o processo global de interpretação dos signos.

No primeiro interpretante, o imediato, estão implicadas as possibilidades potenciais ainda não realizadas, possibilidades de interpretação internas ao signo, ou seja, é a primeira impressão do signo. É o que Peirce chama de nível de primeiridade.

No interpretante dinâmico encontra-se uma ação mais complexa, uma vez que a ação ocorre fora do signo, mesmo sendo uma ação influenciada por ele, ou seja, é um efeito produzido sobre um determinado intérprete, numa dada situação, num dado estado. Nesse sentido, “o interpretante Dinâmico é a determinação de um campo de representação exterior ao signo (sendo tal campo a consciência de um intérprete), determinação esta que é afetada pelo signo” e é considerada como secundidade (SANTAELLA, 1995, p.98).

Com efeito, pode-se dizer que este interpretante corresponde ao que Peirce chama de significado psicológico do signo, efetuado diretamente num processo comunicativo, na mente do intérprete, no momento em que se une a outros signos interpretados, formando uma cadeia semiótica, partindo daí ao início da consolidação de um significado (RANSDELL apud SANTAELLA, 1995).

Finalmente, com o interpretante final, se estabelece a maneira pela qual o signo está relacionado diretamente ao seu objeto. Embora a palavra final denote o encerramento do “processo de semiose”, aqui ela é empregada como um limite ideal, aproximável, servindo de base para que o interpretante dinâmico se ajuste e desenvolva realmente a significância através de um novo signo.

Umberto Eco (2007, p.98), abordando o pensamento peirceano, traduz esse processo afirmando que para se estabelecer o significado de um signo é necessário “substituí-lo por um outro signo ou conjunto de signos, que por sua vez é interpretável por um outro signo ou conjunto de signos e assim por diante *ad infinitum*”. Para o autor, Peirce define signo como “cada coisa que induz outra coisa (o seu interpretante)”. Complementa afirmando que o interpretante final pode consistir no sentido profundo e no efeito conclusivo de um texto, aspecto que o próprio Peirce considera quase impossível de se concretizar devido aos problemas de tradução no nível intermediário, tal como afirmado acima por Santaella.

Esse processo de representação é uma relação denominada por Peirce de “genuinamente triádica”, com papéis definidos distintamente para os três elementos (o signo, o objeto e o interpretante), mas com uma interdependência tal que chegam praticamente a ocupar um mesmo espaço dentro do processo, já que um inexistiria sem o outro. No entanto, do interpretante, no processo, espera-se uma natureza de lei, de regra, por intermédio do qual o signo é desenvolvido e transformado em outro signo, num processo de auto geração, estabelecendo uma relação clara e objetiva com o seu objeto. Peirce explica que é possível esclarecer o sentido real de um signo apontando um objeto,

fazendo uma mímica ou até mesmo lendo um discurso completo; são atitudes que auxiliam a entender o signo externado.

Partindo desses princípios, verifica-se que o interpretante se encontra no processo como elemento mediador entre o signo e o objeto, mas que depende principalmente do signo para que alcance uma interpretação correta do objeto, seja este objeto um acontecimento, um pensamento, uma coisa, etc. A formalização das relações e as definições em si podem ser consideradas interpretantes, cuja formulação vem de um processo pragmático.

Isto posto, pode-se concluir que a representação em vários contextos caracteriza uma função sónica ou um processo de utilização sónica, uma vez que constituiu um processo de apresentação de algo por meio de signos ou símbolos. Adentrando na Linguística, e considerando os elementos acima, os signos que representam uma palavra, um conteúdo ou um objeto, devem levar em consideração as características sintáticas e semânticas da linguagem e do texto em que esta palavra ou conteúdo está inserido para que haja uma interpretação correta por parte do intérprete. (EDMUNDSON apud KOBASHI, 1996, p.07). Assim, o elemento primeiro para uma boa representação é o contexto, levando em consideração a totalidade da cultura, do econômico e do social que o envolvem; são estes elementos que constituem o interpretante, o responsável pela produção de sentido, na abordagem peirciana.

Eco, refletindo sobre a supremacia do interpretante para a produção de sentidos de um signo, acredita ser o interpretante o responsável pela máxima pragmática de Peirce.

Consideremos quais os efeitos que poderiam, concebivelmente, ter as consequências práticas que nós concebemos que os objetos das nossas concepções têm. Então, nossa concepção desses efeitos é a totalidade da nossa concepção do objeto". (PEIRCE apud ECO, 2007, p.99)

Olhando pelo aspecto pragmático do processo, é possível estabelecer relação direta entre o signo e seu interpretante: quando se pensa sobre um

signo, automática e involuntariamente, emerge o conhecimento prévio acerca dele. Em outras palavras, para um leitor assíduo de ficção, é fácil diferenciar uma obra de Shakespeare de um clássico policial de Alan Poe, pois são elementos que fazem parte de sua cultura (ECO, 2007).

No entanto, Eco explica que para Peirce essas respostas automáticas e involuntárias só podem ser consideradas como interpretantes se forem do tipo “que me faz saber algo mais”. Assim, para falar de uma obra de Shakespeare é importante, numa representação, esclarecer ao menos se esta se encontra na categoria de romance ou de teatro para que o caminho do processo se desenvolva corretamente. Portanto, não basta substituir, é necessário mostrar algo mais para precisar o significado. Não basta, por isso, dispor um signo para representar, é necessário trabalhar conceitos e relações de contexto. E o algo mais precisa estar em consonância com a realidade de quem recebe o signo, num processo pragmático coerente com o contexto.

Essas nuances consideradas essenciais no processo de interpretação, nada mais são do que elementos da vida cotidiana que fazem de um receptor de mensagens um intérprete multicultural. É nesse sentido que as teorias Semióticas e Linguísticas favorecem significativamente a Documentação, que necessita de quadros teóricos para compreender os processos relacionados à passagem do texto original para uma representação através de linguagens de representação.

Quando considera o contexto, incluindo aí a realidade cultural e o pensamento humano como elemento mediador da produção de significados através do interpretante, Peirce coloca o pragmatismo como questão central das discussões filosóficas sobre a linguagem (PEIRCE, 1977 apud SOUZA, 2010).

2.2 Pragmática Linguística e informação contextual

A abordagem pragmática das atividades de tratamento e organização da informação se preocupa em ressaltar o sentido de certas formas linguísticas, sentido determinável somente no uso da língua, dentro de um contexto social. Esse processo envolve fenômenos discursivos, comunicativos e sociais, portanto, dentro de regras compartilhadas, caracterizando assim a preocupação com a relação entre os signos e seus intérpretes, os usuários dos signos (MORRIS apud ARMENGAUD, 2006).

Para Morris (1976), a diversidade de signos usados por um mesmo intérprete resulta em inúmeras possibilidades de relações sintáticas. Em outras palavras, um signo que possa ser usado como apresentação de vários objetos, ou vários signos que possam ser usados para a apresentação de um único objeto, são relações que têm respostas no modo de uso dos termos.

Portanto, a pragmática favorece as atividades que envolvem operações de organização de registros documentais em sistemas de informações porque considera os contextos em que ocorrem as referidas operações.

O termo “pragmática” foi cunhado em 1938, por Charles Morris, que o definiu como o estudo da linguagem em relação aos seus falantes, ou usuários. Essa perspectiva foi marcante no início do século XX, quando a semântica e a sintaxe não abordavam suficientemente as nuances da linguagem natural. Nessa época, ancorados no positivismo lógico, os estudos semânticos e da sintaxe eram voltados para a linguagem formal da lógica, sendo a linguagem natural vista como imperfeita para tal atividade, ou seja, os filósofos da época a julgavam sem valor para a atividade científica (IBAÑOS e SILVEIRA, 2002).

Nessa perspectiva, Morris parte de consagradas definições das três dimensões para justificar a importância da pragmática como complemento dos estudos lingüísticos:

Sintaxe é o estudo das relações sintáticas dos signos entre si, abstraído das relações dos signos com os objetos ou com os intérpretes(...) a semântica trata das relações dos signos com seus “designata” e também com os objetos que eles podem denotar ou realmente denotam(...) e a pragmática designa a ciência da relação dos signos com os seus intérpretes. (MORRIS, 1976, p.27-50)

Corroborando Morris, Ibaños e Silveira (2002), em suas reflexões sobre as interfaces da semântica e da pragmática, acreditam que essas disciplinas científicas, embora distintas, são dependentes metodologicamente. Isso porque, enquanto a semântica visa apreender as condições de verdade das sentenças, a pragmática busca as condições do uso apropriado das mesmas, ambas estabelecendo o caminho para um significado aproximado das sentenças. Existem, portanto, significados que as sentenças não trazem explicitas em si, necessitando dos contextos em que estão inseridas, para serem interpretadas.

2.3 Abordagens pragmáticas na C.I.

A virada pragmática vem influenciando a CI. Dentre as abordagens pragmáticas podem-se destacar Cardoso (1994), Araújo (2003), Capurro (2003), Gonzáles de Gomez (2007), Kobashi (2010), entre outros, que olham a informação em contextos, descartando a visão puramente formalista construída nos primórdios da CI.

Araújo, que integra a CI às Ciências Sociais, vai ao encontro das teorias de Borko e Buckland:

Ao discutir a realidade como algo que é constituído socialmente e não com uma existência em si mesma, independente dos sujeitos que conhecem, [abre-se] caminho para uma compreensão da informação não como um dado, uma coisa que teria um significado e uma importância per se, mas como um processo, como algo que vai ser percebido e compreendido de variadas formas de acordo com os sujeitos que estão em

relação – o que vai na contramão tanto da definição de Borko (1968) sobre o comportamento e o fluxo da informação excluindo os sujeitos, quanto da definição de Buckland (1991), que vê “a informação como coisa”. A questão da intersubjetividade conformada a partir da informação se torna central para a compreensão dos diferentes planos de realidade, da distinção entre as diferentes formas de conhecimento e dos mecanismos de sua configuração e legitimação. Os sujeitos precisam necessariamente ser incluídos nos estudos sobre informação e, sobretudo precisam ser incluídos em suas interações cotidianas, formas de expressão e linguagem, ritos e processos sociais. (ARAÚJO, 2003, p.25).

Reforçando essa afirmativa, Cardoso traz três elementos que apóiam as atividades informacionais na área e sinalizam para uma perspectiva social e pragmática (CARDOSO, 1994, p. 111 e 112):

- historicidade dos sujeitos e dos seus respectivos objetos, colocando-os em uma relação culturalmente determinada;

- a sociedade é constituída de uma estrutura orgânica em que o conjunto dá o sentido necessário aos fenômenos, sendo quase impossível se estudar um fenômeno social isolado e encontrar respostas condizentes com a realidade;

- tensionalidade presente na sociedade, aquele que lapida a disputa pelo poder através da cultura, da política, da economia, elementos sociais por natureza.

Refletindo sobre as propostas dos autores acima citados, podem-se inferir os elementos que trazem à tona a natureza social e cultural da informação, incorporando irreversivelmente o homem, enquanto ser social e cultural (responsável pela formação de significados e sentidos das mensagens que formulam a comunicação entre os indivíduos), no centro das pesquisas em CI e, conseqüentemente, nos trabalhos de organização e tratamento da informação.

A ligação entre sujeito e informação só se efetua se o primeiro consegue traduzir ou interpretar adequadamente a informação recebida. Assim, é

necessário contextualizar a informação no processo de representação, levando em consideração as características culturais e sociais do sujeito, bem como a relação que essas características têm com outros mundos, outras culturas.

Os parâmetros ou diretrizes são indispensáveis para os trabalhos de representação. Através das diretrizes, ou regras, tem-se a possibilidade de minimizar as inconsistências entre aquilo que a sociedade demanda informacionalmente e aquilo que é oferecido pelos sistemas.

As inconsistências e dificuldades de manuseio da Informação são elementos de discussão na área da CI, incluindo suas múltiplas interpretações, tópico que aumenta a cada dia sua extensão conceitual. A discussão de González de Gómez sobre o termo traduz o problema:

(...) ‘informação’, como construção e sintoma da sociedade contemporânea, designa algo que se desdobra em vários planos de realização — todos eles, quaisquer que sejam suas características, transcendidos por valores semânticos e quadros normativos, da ordem do ideacional, e ancorados no solo de sua inscrição ou manifestação, da ordem do visível e do temporal (GONZALEZ DE GÓMEZ, 2007, p.03).

Na área da CI esses aspectos, quando levados em consideração, contribuem para constituir o chamado perfil do usuário, aspecto importante do processo de significação. Quando se dá a significação ou a formação de sentidos, esclarecendo “qual é o caso em que a informação é o caso”, acontece o que González de Gómez (2007, p.4) chama de “ação Informativa”. É o momento em que há assimilação da mensagem ou informação emitida, quando o conteúdo da mensagem é consoante com a realidade do usuário.

Para Gonzáles de Gómez (2007), as qualidades e as configurações em que se projeta a ação informativa são variadas e só podem ocorrer dentro de contextos específicos, sendo, portanto, uma ação essencialmente pragmática. Existem, portanto, elementos *a priori* que regulam e condicionam a aceitação de algo ou coisa como Informação.

Neste sentido, Kobashi explica que para haver um conhecimento prévio das informações de um sistema de informação pelo usuário é necessário dispor de uma linguagem de representação adequada, que leve em consideração esses elementos *a priori* e estabeleça “uma intercomunicação entre sistema e usuário”. Essa linguagem, denominada documentária, é pré-estabelecida a partir de aspectos contextuais, sociais e culturais. Elas não são, pois, “meras nomenclaturas ou listas de palavras e expressões utilizadas para etiquetar documentos para armazenamento. Ao contrário, são instrumentos essenciais para haver interação e diálogo entre sistemas de informação e usuários” (Kobashi, 2007³). Nesse sentido, o contexto se torna indispensável para os trabalhos da CI, que busca obter índices crescentes de eficácia nos resultados da recuperação da informação.

Como se pode perceber, o contexto é uma constante em todo o processo de significação da linguagem. Assim, subentende-se que deve estar no centro desse processo o sistema simbólico utilizado pelo sujeito, ou seja, a linguagem derivada de seus ritos sociais, como argumenta Araújo (2003). Nesse cenário, o contexto se apresenta como um conceito essencial, que esclarece a “situação concreta em que os atos de fala são emitidos, ou proferidos, o lugar, o tempo, a identidade dos falantes etc., tudo que é preciso saber para entender e avaliar o que é dito” (ARMENGAUD, 2006, p.13). Assim, a importância do contexto como instrumento para a organização da informação é indiscutível. Quando as informações são abstraídas do lugar, do tempo, da identidade dos documentos e dos seus usuários, podem se tornar informações ambíguas destituídas de significado.

Para Armengaud (2006), saber da importância do contexto para os atos de fala é mais fácil do que saber delimitá-lo. É uma importante ressalva para o tratamento da informação porque é possível, a partir dessa delimitação, nortear o uso eficaz do contexto nas atividades informacionais, bem como verificar a própria necessidade de uso, pois para algumas linhas de raciocínio o contexto nem sempre é necessário na validação de uma sentença (BAR-HILLEL, 1982; IBÁÑOS E SILVEIRA, 2002). “O gelo flutua sobre a água” e “A neve é branca”

³ Documento eletrônico

são exemplos clássicos apresentados como verdades absolutas, perfeitamente entendidas e avaliadas, sem precisarem necessariamente de informações prévias para sua elucidação.

Armengaud (2006, p.79) estabeleceu 3 tipos de “contextos”, cada um deles com diferentes constituintes:

- a) **Contexto circunstancial, factual, existencial, referencial** – identidade dos interlocutores, seu ambiente físico, o lugar e o tempo em que suas sentenças são expressas. Tudo o que faz parte do estudo dos indexais.
- b) **Contexto situacional ou paradigmático** – culturalmente mediado. “Situação” qualificada e socialmente reconhecida, comportando uma ou várias finalidades e um sentido imanente partilhado pelos protagonistas pertencentes a uma mesma cultura. As práticas discursivas se inserem em situações definidas às vezes tacitamente, às vezes por proclamação específica. As sentenças nelas proferidas fazem sentido; transplantadas para outra situação, deixam de fazer sentido e parecem incongruentes.
- c) **Contexto interacional** – encadeamento dos atos da fala em uma sequência interdiscursiva. Os interlocutores desempenham papéis propriamente pragmáticos: propor, objetar, retratar. Um ato de fala chama outro, determinado por pressão sequencial.
- d) **Contexto pressuposicional** – constituído por tudo o que é igualmente presumido pelos interlocutores, suas pressuposições e/ou crenças, suas expectativas e intenções.

Os contextos acima, aplicados ao contexto informacional, auxiliam a perceber a pertinência cultural, ou seja, os termos adequados para cada lugar e identidade, o que promove a interação entre o usuário e a informação tratada (ou signos carregados de interpretações), respeitando as pressuposições comuns por estarem imersas na mesma realidade cultural.

No ambiente da organização e tratamento da informação, os contextos propostos por Armengaud podem ser sistematizados por meio de métodos, procedimentos e regras adequadamente adaptadas às condições de produção e recepção em sistemas de informação.

- a) **Contexto circunstancial, factual, existencial, referencial/ Dispositivos informacionais institucionalizados:** consideram-se aqui as diferentes unidades e dispositivos de informação (serviços de informação, bibliotecas, arquivos e museus, sejam eles tradicionais, eletrônicos ou digitais) e sua vinculação institucional.
- b) **Contexto situacional ou paradigmático/Áreas de conhecimento/ atividades:** contexto no qual se explicitam as finalidades da unidade de informação; definem-se, assim, as áreas do conhecimento que integram os dispositivos informacionais e a linguagem do sistema.
- c) **Contexto interacional/Interfaces de mediação: contexto de organização e busca de informação.** Definem-se neste âmbito as hipóteses de organização e de acesso à informação. É, portanto, instância dialógica de negociação de sentidos. As mediações podem ser interpessoais ou realizadas por meio de interfaces eletrônicas.
- d) **Contexto pressuposicional/Expectativas: contexto ao qual se subordinam as Interfaces de mediação.** As expectativas são compartilhadas tacitamente ou por proclamação. São antecipações determinadas pelo Contexto circunstancial e situacional.

O contexto, em diferentes níveis, é elemento presente nas ações semânticas e pragmáticas próprias do tratamento da informação. Desse modo, considerar a linguagem na sua dimensão social e cultural é entender a substancialidade e a validade da pragmática linguística para a área. Sua discussão implica a necessidade de adoção da perspectiva pragmática no contexto da representação e da recuperação da informação.

Assim, para atender com eficácia às demandas de informação, no cenário atual, é necessário substituir os métodos anteriores de tratamento da informação por modelos pragmáticos que comportem contextos, de forma a unir termos de representação e realidade social de um determinado domínio do conhecimento.

3 SOBRE ONTOLOGIAS

3.1 Definições e Histórico

O termo Ontologia tem origem na filosofia clássica grega, com a seguinte definição: ontos=ser e logos=palavra (ALMEIDA, 2003,p.7). Porém, os estudos ontológicos tiveram início na época de Aristóteles, quando ainda se intitulavam estudos Metafísicos. Compreendia “o conhecimento das coisas divinas, ao mesmo tempo que os dos princípios das ciências e da ação” (LALANDE, 1993, p.666). Na modernidade, o conceito de Metafísica foi ampliado, comportando ora o conhecimento de certos seres, ora uma certa ordem da realidade.

Na área computacional, o termo foi apropriado, na década de 1990, pelos engenheiros de software da área da Inteligência Artificial (IA). Caracterizaram a Ontologia como “um relato sistemático da Existência. Para sistemas baseados em conhecimento, o que ‘existe’ é exatamente o que pode ser representado” (GRUBER, 1993, p.200). Dessa forma, o termo foi empregado para designar um novo aparato tecnológico útil para a representação do conhecimento na Web semântica. No entanto, para adaptá-lo aos trabalhos de representação, o conceito foi simplificado passando a designar não o “Ser”, mas um dado “domínio” de conhecimento, que envolve conteúdos e processos específicos, dando conta não apenas do que existe, mas também de toda estrutura do que existe. Nesse sentido, Sowa (1999)⁴ afirma que “a intenção da Ontologia é estudar as categorias de coisas que existem ou podem existir em algum domínio”. Nesse caso, pode-se acrescentar à reflexão de Sowa que a ontologia a que se refere está mais direcionada às ontologias de domínio, que abordam áreas específicas de conhecimentos.

⁴ Documento eletrônico

Na Ciência da informação, com as exigências de soluções cada vez mais eficazes para o tratamento da documentação digital, as Ontologias surgiram como um novo tipo de ferramenta para otimizar o tratamento e a recuperação de informação. Segundo Campos (2009), o termo surgiu na área entre 2003 e 2004, quando da criação da Web 2.0, ambiente que supõe a interação do usuário com o processo de inserção de conteúdos. Dessa forma, as Ontologias são vistas como “mediações de linguagens adotadas em sistemas de informação, compostas por estruturas informacionais contextualizadas, derivadas de esquemas intelectuais mais complexos” (CAMPOS, 2009, p.199). São vocabulários (uma denominação herdada dos conhecidos tesouros) relacionados de maneira sistemática que permitem que os sistemas automatizados interpretem conceitos.

Na verdade, o conceito do termo Ontologia, no contexto da Web, não tem um significado estabilizado. Guarino (1998, p.4) define Ontologia como “uma maneira de se conceitualizar de forma explícita e formal os conceitos e restrições relacionados a um domínio de interesse”, ou ainda, “um conjunto de axiomas lógicos concebidos para explicar o significado pretendido de um vocabulário”. Ela é, portanto, um instrumento para representar conceitos de uma determinada área do conhecimento, de maneira objetivada e eficaz, não só para a leitura do homem, mas também da máquina.

Gruber (1993, p.2), por sua vez, define Ontologia como “uma especificação explícita de uma conceitualização”, caracterizando o termo conceitualização como um modelo de entidades, relações, axiomas e regras de algum domínio específico. O autor complementa salientando que a importância desse instrumento está calcada na natureza formal e compartilhada, ou seja, ela deve ser reconhecível por máquinas e disseminada entre sistemas e grupos sociais. Embora seja um conceito bastante citado, ele foi objeto de muitas críticas. Assim, prosseguindo em suas reflexões, Gruber elaborou um novo conceito, mais amplo e mais aceito. Nessa nova definição, ele detalha não só a natureza conceitual do termo, mas também a estrutura e suas respectivas relações para representar conhecimento:

Ontologias são acordos sobre conceitualizações compartilhadas. Conceituações compartilhadas incluem quadros conceituais para o conhecimento do domínio de modelagem, protocolos de conteúdo específico para a comunicação entre os agentes operacionais inter e acordos sobre a representação das teorias de domínio particular. No contexto de partilha de conhecimento, as Ontologias são especificadas na forma de definições do vocabulário de representação. Um exemplo muito simples seria um tipo de hierarquia, especificando as classes e suas relações de subsunção. Esquemas de Banco de dados relacional também servem como Ontologias, especificando as relações que podem existir em algum banco de dados compartilhado e as restrições de integridade que devem ter para eles. (GRUBER, 1994 apud GUARINO, 1997 a)

Nessa abordagem, Gruber deixa claro que a conceitualização não é a Ontologia em si, mas são acordos que a formalizam, acordos em que podem existir vários componentes que operam na sistematização e formalização dos conceitos, dependendo da finalidade da Ontologia. Ou seja, as Ontologias definem as normas que regulamentam as combinações entre os termos e suas relações.

O aspecto formal e semântico faz das Ontologias a principal ferramenta da Web semântica para capacitar os sistemas e programas a entender e processar dados disponíveis no ambiente hipermídia. Esse processo requer padrões e ferramentas para atribuir significado às sentenças e aos enunciados inseridos na Web, bem como compartilhar de forma prática os conteúdos (Ribeiro, 2008).

Corroborando esta reflexão, Lima e Carvalho (2005, p.3) afirmam que a relevância das Ontologias para a Web deve ainda ser tratada sob três aspectos:

- **Identificação de contexto:** por exemplo, quando dois agentes de *softwares* trocam informações sobre o termo *braço* é preciso assegurar em que contexto

esse termo está sendo referenciado, haja vista que o termo é usado em várias áreas, como medicina, movelaria, oceanografia, entre outros.

- **Fornecimento de definições compartilhadas:** por exemplo, se uma aplicação X possui uma Ontologia que define uma loja que vende carros e uma aplicação Y possui outra Ontologia que define uma loja que vende veículos, logo se percebe o problema caso ambas queiram intercambiar informações. Este problema é facilmente compreendido pelo ser humano, mas não por máquinas. Este tipo de problema pode ser resolvido se as Ontologias proverem relações de equivalência, ou seja, se uma ou as duas Ontologias possuírem informações dizendo que o termo carro da aplicação X é equivalente ao termo veículo da aplicação Y.

- **Reúso de Ontologias:** por exemplo, se uma determinada pessoa X já tem construída uma Ontologia (sobre medicina, por exemplo) que define um conjunto de termos que outra pessoa Y também necessita, não há porque a pessoa Y crie outra Ontologia, isto é, refaça algo já feito. Pode-se simplesmente fazer uso da Ontologia já criada.

Itens como esses devem ser considerados na implementação e uso de Ontologias, pois eles permitem padronizar vocabulários e compartilhar dados de diferentes dispositivos.

3.2 Elementos de Composição

Mesmo sendo ferramentas que apresentam uma variedade significativa de formas e estruturas, existem alguns critérios que Gruber (1995) considera importantes para a implementação de qualquer tipo de Ontologia:

- **Clareza e Objetividade:** uma Ontologia deve apresentar de forma clara e objetiva as definições dos termos, assim como uma documentação. Aconselha-se que uma definição seja declarada por meio de axiomas lógicos e as definições devem ser documentadas em linguagem natural;

- **Coerência:** uma Ontologia deve ser coerente, o que significa permitir inferências consistentes com base nas definições. A definição dos axiomas deve ser lógica e a coerência deve atingir também os conceitos definidos informalmente, que devem ser expressos em linguagem natural e conter exemplos. Se um conceito for inferido a partir de axiomas e contradigam uma definição ou exemplo dado informalmente, então a Ontologia é incoerente e deverá ser revista;

- **Extensibilidade:** uma Ontologia deve permitir que novos termos possam ser definidos para usos especiais, baseados nos vocabulários existentes, de maneira que não seja requerida a revisão das definições previamente existentes.

Além de critérios em comum, Gruber considera também que alguns elementos são básicos nos vários tipos de Ontologias. Esses elementos são (GRUBER, 1996 apud ALMEIDA e BAX, 2003, p.9):

- **Classes** (organizadas em uma taxonomia) – são conceitos que representam qualquer coisa do domínio, incluindo objetos do domínio, a descrição de uma tarefa, de uma função ou de uma ação, entre outras. Nesses conceitos estão embutidas informações associadas ao domínio trabalhado, denominadas atributos. Correspondem às:

- classes dos modelos orientados a objetos
- entidades do modelo relacional
- termos do modelo lógico

- **Relações** – representam os tipos de interações entre os conceitos do domínio, ou seja, expressam a natureza das ligações entre dois conceitos. São definidas formalmente como qualquer subconjunto de um produto de conjuntos. Correspondem às:

- associações, agregações e atributos dos modelos cujos valores são objetos
- relações do modelo relacional
- predicados do modelo lógico

- **Funções** – São relacionamentos especiais entre elementos. Elas designam os vários elementos envolvidos nesse tipo de relação, dependentes uns dos outros, sendo porém cada um único em relação a outro.

- **Axiomas** – usados para modelar e especificar sentenças consideradas sempre verdadeiras em uma Ontologia, classificadas como estruturais e não estruturais. Os axiomas especificam também as restrições sobre as interpretações de tais sentenças.

- **Instâncias** – utilizadas para representar elementos específicos, ou seja, definição de entidades e relações específicas. Correspondem aos fatos dos sistemas especialistas e programação em lógica, considerados elementos da Ontologia. As instâncias são os conceitos e relações criados na Ontologia.

No âmbito do consórcio de Tim Berners-Lee, as Ontologias devem ser constituídas por apenas três elementos básicos: **classes** ou coisas, nos vários domínios de interesses; **relacionamentos** entre essas coisas; e **propriedades** ou atributos, que essas coisas devem possuir (BREITMAN, 2005, p.31)

A partir desses componentes básicos são construídos diferentes níveis de Ontologias - níveis que dizem respeito à classificação em relação a uma tarefa específica ou um ponto de vista conceitual que direciona a representação de um objeto, ou seja, a abrangência e o domínio do conhecimento que se quer representar.

3.3 Tipos de Ontologias

A tipologia mais comumente encontrada na literatura é a sugerida por Guarino, que classificou as ontologias, de modo geral, nas seguintes duas formas: “com base no grau de detalhe utilizado para caracterizar uma conceituação ou em função do objeto de conceituação”. Nesse sentido, o autor (Guarino, 1997b, 1998) redefine alguns níveis de classificação apresentados por van Heijst, Schreiber e Wielinga (1996):

- a) **Ontologias de alto nível** – descrevem conceitos muito gerais como espaço, tempo, matéria, objeto, acontecimento, ação, etc, que são independentes de um problema particular ou domínio. Elas são consideradas adequadas para aplicação em sistemas que atendem comunidades não específicas de usuários.
- b) **Ontologias de domínio e Ontologias de tarefas** – descrevem, respectivamente, o vocabulário relacionado a um domínio genérico (como a medicina, indústria automobilística, computação) ou uma tarefa genérica ou atividade (como diagnóstico ou vendas), especializando os termos introduzidos no nível superior da Ontologia.
- c) **Ontologias de aplicação** – descrevem conceitos dependentes tanto de um domínio particular como de uma tarefa específica, que muitas vezes são especializações de ambas as Ontologias relacionadas. Esses conceitos correspondem, via de regra, aos papéis desempenhados por entidades de domínio ao executar uma determinada atividade.
- d) **Ontologias não-refinadas** – têm um número mínimo de axiomas e seu objetivo é ser utilizado por usuários que compartilham uma determinada visão de mundo. Acredita-se que esse tipo de Ontologia deva ser usada *off-line* e somente para referência.
- e) **Ontologias refinadas** – requerem linguagem de alta expressividade e têm um grande número de axiomas. As Ontologias não-refinadas têm mais chance de ser compartilhadas e deveriam ser usadas *on-line* para dar suporte à funcionalidade dos sistemas de informação.

Em estudos mais recentes, Almeida e Bax (2003) relacionam diversos tipos de Ontologias que seguem de perto a tipologia de Guarino (1997, 1998), acrescida de uma Ontologia, a de representação, conforme descrição abaixo:

- **Ontologias genéricas:** referem-se a conceitos, de uma forma geral, os quais independem de um domínio particular. Por exemplo, espaço, tempo, evento.
- **Ontologias de domínio:** apresentam conceitos sobre um determinado domínio, utilizando um vocabulário a partir de um domínio mais genérico. Por exemplo, Ontologias de veículos, documentos, computação, medicina.

- **Ontologias de tarefas:** tratam de conceituações para resolução de problemas, sem levar em consideração os domínios nos quais ocorrem. Delineiam o vocabulário de acordo com uma atividade ou tarefa genérica, por exemplo, a de realizar um diagnóstico, ou vender roupas.
- **Ontologias de aplicação:** relaciona-se aos conceitos inerentes a um domínio e tarefa específica, como por exemplo, uma Ontologia para aplicação no setor de carros de luxo.
- **Ontologias de representação:** explicam as conceituações que fundamentam os formalismos de representação de conhecimentos, como a Ontologia de *frames*, utilizada em ontolingua.

Para Wache (2001) quase todas as abordagens são utilizadas para a descrição explícita da informação semântica de origem, mas existe uma diferença na maneira estratégica de construção de cada abordagem, dentre as quais algumas oferecem subsídios para integração de diferentes ontologias. Para o autor, são três as principais abordagens de integração a serem consideradas para as ontologias, principalmente para as de domínio:

Abordagem única: é a mais simples, que pode ser simulada por outras. Usa uma Ontologia global que fornece um vocabulário comum e compartilhado para a especificação da semântica. Nessa arquitetura, todas as fontes de informações estão relacionadas a essa única Ontologia global, desenvolvida independentemente de outras fontes e de suas Ontologias. Para isso, devem ser descritos para esse sistema um modelo independente de cada fonte de informação, relacionando os objetos dessas fontes para o modelo de domínio global, bem como descrever as relações para esclarecer a semântica dos objetos de origem e ajudar a encontrar outros objetos semanticamente correspondentes, fazendo um mapeamento inter-Ontologias (WACHE, 2001).

O problema dessa abordagem é que embora ofereça um vocabulário compartilhado que permite combinar Ontologias especializadas, por meio da importação de módulos de outras Ontologias – combinação essa suportada por formalismos de representação –, se esses novos módulos não tiverem uma

visão conceitual comum ao domínio da Ontologia global que os comporta, a tarefa se torna quase impossível, pois ela não permite o desdobramento conceitual. Ou seja, as mudanças nas fontes de informações afetam diretamente a conceituação do domínio representado.

Figura 2: Abordagem ontológica Simples
Fonte: Wache, 2001

Abordagem de múltiplas Ontologias: nessa arquitetura, as desvantagens da abordagem simples são superadas, uma vez que se podem modificar as fontes de informações, bem como adicionar e remover as fontes no decorrer do processo ou uso. Assim, percebe-se que essa abordagem não tem compromisso com um vocabulário comum sobre a Ontologia global, o que facilita o trabalho de construção.

No entanto, se por um lado essa flexibilidade facilita a construção, por outro, a falta de um vocabulário comum torna muito difícil a comparação e a integração semântica de diferentes fontes e Ontologias. Para solucionar esse problema é necessário agregar um formalismo de representação de inter-mapeamento que identifica termos semanticamente correspondentes em fontes e Ontologias diferentes. Um exemplo desse formalismo é o KRAFT, um sistema em que as traduções entre diferentes Ontologias são feitas por agentes mediadores especiais que podem personalizar as traduções entre diferentes Ontologias ou mesmo entre diferentes línguas.

Figura 3: Abordagem Ontológica Múltipla
Fonte: Wache, 2001

Abordagem híbrida: foi desenvolvida para resolver os problemas existentes nas duas abordagens citadas anteriormente. Nela, a semântica de cada fonte é descrita por suas próprias Ontologias, como na abordagem múltipla. Mas, para que haja a comparação entre as diferentes Ontologias integradas, elas são construídas sobre um vocabulário comum compartilhado, em que são utilizados termos básicos, chamados de primitivas de um domínio. Para a criação de termos mais complexos são feitas algumas operações sobre as primitivas. Desse modo, os termos baseados em uma primitiva, tornam-se mais fáceis de serem comparados do que na Ontologia múltipla.

A vantagem dessa abordagem é que novas fontes podem ser adicionadas com facilidade, sem a necessidade de modificações do mapeamento ou no vocabulário controlado. Os pontos fracos da abordagem híbrida consistem na existência de Ontologias que não podem ser facilmente reutilizadas, mas devem ser redesenvolvidas, pois todas as fontes ontológicas devem referenciar o vocabulário compartilhado.

Figura 4: Abordagem Ontológica Híbrida
Fonte: Wache, 2001

No quadro a seguir, de Wache (2001), podem ser percebidas com maior clareza as diferenças entre as três abordagens.

Quadro 1: Vantagens e desvantagens da Ontologia diferentes baseadas abordagens de integração

	Abordagens Ontológicas simples	Abordagens Ontológicas múltiplas	Abordagens Ontológicas híbridas
Esforço de implementação	Fáceis	Dispendiosos	Razoáveis
Heterogeneidade semântica	Suportam pontos de vista homogêneos de domínio	Suportam pontos de vista heterogêneos	Suportam pontos de vista heterogêneos
Adição / remoção de fontes	Necessidade de alguma adaptação na Ontologia global	proporcionam uma nova fonte de Ontologia; Relativo à outras Ontologias	Proporcionam uma nova fonte de Ontologia
Comparação de Ontologias	————	Difíceis pela falta de um vocabulário comum	Simples porque usam um vocabulário comum

Fonte: Wache, 2001

Essas abordagens têm sido concretizadas por meio de metodologias específicas que incluem uma linguagem lógica formal, na maioria das vezes

baseada em cálculo de predicados, bem como em ferramentas tecnológicas.

3.4 Metodologias de construção

Assim como existem vários tipos e abordagens de Ontologias, existem várias metodologias utilizadas em sua construção. Apresentamos, a seguir, as principais metodologias citadas por eles:

Ontologias de domínio

- 1) **Metodologia GRUNINGER e FOX:** derivada da experiência de desenvolvimento de Ontologias para domínios corporativos, é muito utilizada para a construção de Ontologias sobre negócios. Foi proposta por Michael Gruninger, em 1995, e Mark Fox em 1992, tendo como base para o seu desenvolvimento a experiência obtida no projeto Toronto virtual Enterprise – conhecido como projeto Tove, cujos princípios teóricos e metodológicos encontram-se na inteligência artificial.
- 2) **Método KACTUS:** aplicado no projeto de mesmo nome, resultou numa Ontologia do domínio de redes elétricas. É um método que dá ênfase à organização de bases de conhecimento que podem ser compartilhadas e reusadas em diferentes sistemas baseados em conhecimento.
- 3) **Metodologia METHONTOLOGY:** foi aplicada na construção de uma Ontologia do domínio da Química. Foi desenvolvida no laboratório de Inteligência Artificial da Universidade Politécnica de Madri entre 1996 e 1997. Contempla um conjunto de estágios de desenvolvimento (especificação, conceitualização, formalização, integração, implementação e manutenção), um ciclo de vida

baseado em evolução de protótipos e técnicas para realizar as atividades de planejamento, desenvolvimento e suporte.

- 4) **Metodologia ON-TO-KNOWLEDGE:** visa a construir Ontologias para aplicações de gestão do conhecimento e são altamente dependentes da aplicação (FERNANDEZ; GÓMEZ-PEREZ; CORCHO, 2004). Auxilia a administração de conceitos em organizações, identificando metas para as ferramentas de gestão do conhecimento e utilizando cenários e contribuições dos provedores ou clientes da organização (ALMEIDA e BAX, 2003).
- 5) **Método 101:** O método foi concebido por Natalya F. Noy e Deborah L. McGuinness, a partir da experiência de desenvolvimento de uma Ontologia de vinhos e alimentos, utilizando o ambiente de edição de Ontologias *Protégé-2000*. Propõe basicamente quatro atividades para o desenvolvimento de uma Ontologia: i) definir classes na Ontologia; ii) organizar as classes em uma taxonomia; iii) definir *slots* (ou propriedades) para as classes e descrever os valores permitidos (denominado *facet*); iv) adicionar valores de *slots* para as instâncias. Tais atividades implicam decisões de modelagem, dentre as quais o método busca enfatizar o processo iterativo de um ciclo de vida de uma Ontologia.

Ontologia de alto nível

- 6) **Metodologia USCHOLD E KING:** foi baseada na prática de construção da Ontologia de alto nível Enterprise. Também utilizada para construção de Ontologias do domínio dos negócios. Foi criada em 1995 por Mike Uschold e Martin King e estendido por Mike Uschold e Michael Gruninger em 1996 na experiência de desenvolvimento da Enterprise Ontology.
- 7) **Método CYC:** considera o conhecimento consensual sobre o mundo e pode ser utilizado na criação de Ontologias para

fundamentar diferentes sistemas inteligentes. Foi criada nos anos 1980 pela Microelectronics and Computer Technology (MCC). É uma ampla base de conhecimento que considera o conhecimento consensual sobre o mundo, incluindo regras e heurísticas para dedução sobre objetos e eventos do cotidiano.

Ontologias de fundamentação e domínio

- 8) Método SENSUS:** baseado na Ontologia de mesmo nome, foi aplicado no desenvolvimento de uma Ontologia do domínio do planejamento de operação militar aérea (VALENTE et *alii*, 1999). Esse método, criado para a construção da Ontologia SENSUS, foi desenvolvido pelo grupo de linguagem natural *Information Sciences Institute* – ISI, para ser usada para fins de processamento de linguagem natural. Propõe alguns processos para estabelecer ligações entre os termos específicos e os termos da Ontologia de alto nível. O resultado de tal processo é uma estrutura de uma nova Ontologia, que é generalizada automaticamente por meio de uma ferramenta denominada OntoSaurus.

Ontologias de tarefa

- 9) Método KBSI IDEF5:** auxilia a criação, modificação e manutenção de Ontologias (JONES; BENCH-CAPON; VISSER, 1998).

Além das metodologias de construção de Ontologias gerais, Almeida e Bax (2003) apresentam outras metodologias para ações mais específicas, como as de construção de Ontologias em grupo, para aprendizado sobre a estrutura de uma Ontologia e para a integração de Ontologias.

Metodologias para construção de Ontologias em grupos

- 1) **Metodologia CO4:** permite discutir o conhecimento introduzido em bases de conhecimento compartilhadas, que correspondem a Ontologias, uma vez que deve haver consenso sobre o conhecimento ali representado. Quando ocorre uma mudança, os usuários são notificados e podem aceitá-la ou não.
- 2) **Metodologia (KA)²:** modela as formas de aquisição do conhecimento usando Ontologias desenvolvidas em conjunto por pessoas de diferentes locais, que utilizam o mesmo padrão; a comunicação e a coordenação são feitas via agentes inteligentes.

Metodologia para aprendizado sobre a estrutura de Ontologias

- 1) **Maedche:** uma Ontologia genérica é convertida em modelo: especificam-se conteúdos e obtêm-se conceitos do domínio a partir das fontes disponíveis; removem-se conceitos genéricos, de forma que apenas os conceitos específicos do domínio permanecem; a estrutura conceitual da Ontologia está estabelecida e obtêm-se as relações existentes entre os conceitos. Novas relações conceituais são induzidas por método de aprendizado.

Metodologias para integração de Ontologias

- 2) **FCA-MERGE:** utiliza métodos de processamento da linguagem natural e análise formal de conceitos sobre as Ontologias que se deseja integrar, resultando em um conjunto de conceitos; o resultado é analisado por

especialistas e transformado em uma Ontologia composta.

- 3) **PROMPT:** algoritmo aplicável a várias plataformas que integram Ontologias, sendo baseada em um modelo de conhecimento genérico; executa algumas tarefas automaticamente (existe participação humana) e identifica inconsistências na Ontologia, sugerindo como removê-las.

Para que as metodologias acima possam ser utilizadas de forma coerente há uma norma que as padroniza. Segundo Silva et al. (2008), a norma, proposta pela organização norte americana National Information Standards Organization – ANSI, em 2005, propõe as linhas gerais para a construção, formatação e manutenção de vocabulários controlados monolíngues, inclusive Ontologias. No entanto, é preciso reconhecer que essa norma, por estar voltada para a construção de instrumentos terminológicos documentários, não é suficiente para dar conta dos vários aspectos das ontologias, em particular, de seus aspectos formais e das ferramentas ontológicas. Os aspectos formais das ontologias serão apresentadas e discutidas na seção 4 desta tese (página 62).

3.5 Ferramentas de Construção de Ontologias

Para a construção de Ontologias são necessários, além de metodologia criteriosa e elementos básicos, as ferramentas ontológicas, que consistem em “mediações de linguagens adotadas em sistemas de informações, compostas por estruturas informacionais contextualizadas, derivadas de esquemas intelectuais mais complexos” (MOURA, 2009, p.199). Vários programas computacionais têm sido desenvolvidos na área da Inteligencia artificial.

As ferramentas mais utilizadas e citadas na literatura atual são: Apollo,

Ontolingua, Ontosaurus, OilEd, OntoEdit, Protégé-2000, WebODE, WebOnto. Para um melhor entendimento dessas ferramentas, apresentamos, a seguir, uma breve descrição de cada uma delas:

Apollo

Fonte: Knowledge Media Institute of Open University (UK)

linguagem de base: OKBC;.

Maiores informações: <http://apollo.open.ac.uk/index.html>

Aplicação desenvolvida por meio de primitivas básicas, como classes, funções, relações, instâncias, entre outras, modeladas de acordo com um sistema de frames. Além disso, a Apollo fornece checagem de consistência na medida em que a Ontologia é desenvolvida. A ferramenta não requer a utilização de uma linguagem específica, podendo ser adaptada a diversos formatos de armazenamento por meio de plug-ins. Essa ferramenta foi desenvolvida em linguagem Java e tem arquitetura aberta (APOLLO, 2003). Aceita as linguagens OCML, Ontolândia, entre outras (LOPES, 2006).

Ontolingua with Chimaera

Ano de lançamento: 2001; versão definitiva em 2002

Fonte: Stanford Knowledge Systems Lab

Lingua base: Ontolingua

Maiores informações: <http://www.ksl.stanford.edu/software/ontolingua/>
<http://www.ksl.stanford.edu/software/chimaera/>

Servidor criado pelo KSL (Knowledge Systems Laboratory – Stanford University), constituído por um conjunto de ferramentas e serviços que suportam a construção de Ontologias compartilháveis entre grupos distribuídos geograficamente. A arquitetura do servidor fornece acesso a bibliotecas de Ontologias, a tradutores de linguagens e a um editor. Editores remotos permitem visualizar e editar as Ontologias e, através delas, aplicações remotas ou locais permitem acessar qualquer Ontologia das bibliotecas, por meio do protocolo OKBC (Open Knowledge Based Connectivity) (CORCHO;

FERNANDEZ-LOPES; GOMEZ-PERES, 2001). O Servidor Ontolingua aceita as linguagens Loom e Ontolingua, entre outras (LOPES, 2006).

Ontosaurus

Fonte: USC Information Sciences Institute, university of Southern California

Linguagem base: Loom

Maiores informações: <http://www.isi.edu/isd/ontosaurus.html>

Constituído por um servidor de navegação de Ontologias, o qual cria dinamicamente páginas em formato HTML para exibição da hierarquia de uma Ontologia. Apresenta dois módulos: um servidor de Ontologias que usa a linguagem LOOM para a representação do conhecimento e um servidor de navegação por Ontologias, que cria páginas HTML de forma dinâmica, incluindo imagens e documentação textual. Apresenta a hierarquia da Ontologia e usa HTML para que o usuário possa editá-la. (SWARTOUT, 1997 apud PAIVA, 2006).

OilEd

Fonte: University of Manchester Information Management Group

Linguagem base: DAML+OIL

Maiores informações: <http://oiled.man.ac.uk/>

Ferramenta para edição de Ontologias baseadas em linguagem OIL e DAML+OIL. Em sua funcionalidade básica, a ferramenta permite definir e descrever classes, slots, entidades e axiomas. As classes são definidas em termos de suas superclasses e restrições de propriedades. Uma inovação apresentada pelo OIEd é a utilização de raciocínio para checagem de consistência entre as classes e para inferência e classificação de relacionamentos. O serviço de raciocínio é provido pelo sistema FaCT, o qual consiste de um classificador de Ontologias via tradução de DAML+OIL para lógica de descrição (BECHHOFFER et al., 2001). A ferramenta foi desenvolvida em linguagem Java e está disponível para uso por meio de licença GPL (General Public License). Essa ferramenta trabalha com as linguagens: RDF, RDF Schema, OIL e DAML+OIL. (LOPES, 2006)

OntoEdit

Fonte: Ontoprise GmbH

Linguagem básica: LF-Logic

Maiores informações: <http://www.ontoprise.de/com/ontoedit.htm>

Ambiente gráfico para edição de Ontologias, que permite inspeção, navegação, codificação e alteração de Ontologias. O modelo conceitual é armazenado usando um modelo de Ontologia que pode ser mapeado em diferentes linguagens de representação. As Ontologias são armazenadas em bancos relacionais e podem ser implementadas em XML, FLogic, RDF(S) e DAML+OIL. Possibilita a exportação das Ontologias criadas para outros formatos como, por exemplo, Document Type Definition (DTD12) (MAEDCHE, 2000 apud ALMEIDA et al., 2005).

Protégé-2000

Fonte: Stanford Medical Informatics

Linguagem base: OKBC model

Maiores informações: <http://protege.stanford.edu/index.html>

Ambiente interativo para projetos de Ontologias, de código aberto (open source), além de ser multi-plataforma, o que a torna a ferramenta mais utilizada atualmente para criação e edição de Ontologias e bases de conhecimentos. Suporta a criação, visualização e manipulação de Ontologias em vários formatos, facilitando a reutilização e o intercâmbio entre elas. Foi desenvolvido em Java pela Universidade de Stanford, a mesma que desenvolveu o servidor da Ontolíngua. Sua arquitetura é modulada e permite a inserção de novos recursos, como plugins desenvolvidos por usuários. Apresenta exemplos que podem ser consultados pelo usuário. Permite gerar saídas gráficas diversas (diagramas de vários tipos). Devido a essas funcionalidades, as aplicações desenvolvidas no Protégé são usadas em resoluções de problemas e tomadas de decisões em um domínio particular.

Ferramenta em constante desenvolvimento, sendo sua versão mais atual, a 3.1 (NOY; FERGERSON; MUSEN, 2000; PAIVA, 2006; LOPES, 2006).

WebODE

Fonte: Technical University of Madrid UPM

Linguagem base: Prolog translation of FOL and frames per OKBC model.

Maiores informações: <http://delicias.dia.fi.upm.es/webODE/>

Ambiente para engenharia ontológica que dá suporte à maioria das atividades de desenvolvimento de Ontologias. A integração com outros sistemas é possível, importando e exportando Ontologias de linguagens de marcação. (ALMEIDA et al., 2005)

WebOnto

Fonte: Knowledge Media Institute of Open University (UK)

Linguagem base: OCML

Maiores informações: <http://kmi.open.ac.uk/projects/webonto/>

Ferramenta desenvolvida pelo KMI (Knowledge Media Institute - Open University). Suporta navegação colaborativa, criação e edição de Ontologias, as quais são representadas na linguagem OCML. Suas principais características são: (a) gerenciamento de Ontologias utilizando interface gráfica; (b) suporte para modelagem de tarefas; (c) verificação de elementos, considerando herança de propriedades e checagem de consistência; (d) suporte ao trabalho colaborativo. WebOnto é um servidor disponível gratuitamente. Possibilita o acesso a mais de 100 Ontologias, as quais podem ser visualizadas sem restrições (CORCHO; FERNANDEZ-LOPES; GOMEZ-PERES, 2001 apud LOPES, 2006).

CIRCA Taxonomy Administrator

Fonte: Applied Semantics, Inc

Base Língua: ProprietaryURL:

http://www.appliedsemantics.com/as_solutions_autocat_taxadmin.shtml

CoGITaNT

Fonte: LIRMM CNRS (France)

Base Língua: CG model

URL: <http://cogitant.sourceforge.net/>

Coherence

Fonte: Unicorn Solutions

Base Língua: XML

URL: <http://www.unicorn.com/pr-overview.htm>

Contextia

Fonte: Modulant

Base Língua: Express

URL: <http://modulant.com/products/>

COPORUM OntoBuilder

Fonte: CognIT AS

Base Língua: RDFS

URL: <http://ontoserver.cognit.no/>

DAG-Edit

Fonte: Berkeley Drosophila Genome Project (BDGP)

Base Língua: Directed (cyclic or acyclic) graph notation

URL: <http://sourceforge.net/projects/geneontology>

DAMLImp (API)

Fonte: AT&T Government Solutions

Base Língua: DAML+OIL

URL: <http://codip.grci.com/Tools/Components.html>

Differential Ontology Editor (DOE)

Fonte: National Audiovisual Institute - INA (France)

Base Língua: XML & CGXM

URL: <http://opales.ina.fr/public/>

Disciple Learning Agent Shell

Fonte: George Mason University , Learning Agents Laboratory

Base Língua: OKBC-like URL: <http://lalab.gmu.edu/>

Domain Ontology Management Environment (DOME)

Fonte: Btexact Technologies

Base Língua: CLASSIC & FaCT

URL: <http://more.btexact.com/projects/ibsr/dome/index.htm>

DUET

Fonte: AT&T Government Solutions

Base Língua: UML

URL: <http://codip.grci.com/Tools/Tools.html>

CODE4

Ferramenta de propósito geral que possui diferentes modos de herança e inferência, uma interface gráfica de fácil uso, um modo de hipertexto para navegação e utilitários para leitura de documentos e gerenciamento léxico. (SKUCE, 1995 apud ALMEIDA et al., 2005)

VOID

Ambiente para navegação, edição e gerenciamento de Ontologias. Por meio de simulações, possibilita o estudo de questões teóricas, como organização de bibliotecas de Ontologias e tradução entre diferentes formalismos (SCHREIBER; TERPSTRA; SISYPHUS, 1995 apud ALMEIDA et al., 2005).

IKARUS (Intelligent Knowledge Acquisition and Retrieval Universal System)

Explora as capacidades cooperativas do ambiente Web. Utiliza representação hierárquica gráfica que permite heranças múltiplas. As declarações que contêm informação são representadas como predicados com sintaxe e semântica definidos ou como fragmentos sem estrutura (SKUCE, 1996 apud ALMEIDA et al., 2005).

GKB-Editor(Generic Knowledge Base Editor)

Ferramenta para navegação e edição de Ontologias por meio de sistemas baseados em frames. Oferece interface gráfica, que permite aos usuários editar diretamente a base de conhecimentos e selecionar a parte de interesse (PALEY; KARP, 1997 apud ALMEIDA et al., 2005).

APECKS (Adaptive Presentation Knowledge Structuring)

Servidor de Ontologias que permite trabalho cooperativo. As Ontologias criadas podem ser comparadas, sendo possível discutir as diferenças e similaridades entre elas (TENNISON; SHADBOLT, 1998 apud ALMEIDA et al., 2005).

Ontomarkup Annotation Tool

Ferramenta baseada em Ontologias incorporar informações semânticas em documentos através de anotações. Contém um componente de marcação que permite a navegação e a marcação de partes relevantes, um componente que aprende regras a partir de exemplos e um componente de extração da informação (VARGAS-VERA et al, 2001 apud ALMEIDA et al., 2005).

Text-to-onto

Ambiente para aprendizado e construção de Ontologias a partir de textos. Os textos pode ser em linguagem natural ou formatados em HTML. O sistema é composto por um módulo de gerenciamento de textos e um extrator de informações. Os resultados são armazenados em XML (MAEDCHE ; VOLZ, 2001 apud ALMEIDA et al., 2005).

4 LINGUAGENS FORMAIS DE CONSTRUÇÃO DE ONTOLOGIAS

A representação do conhecimento em ambiente digital é uma tarefa bastante singular que exige regras, símbolos e linguagens específicas; distingue-se, pois, da representação por meio de linguagens documentárias tradicionais (LDs). A representação em ambiente multimídia requer atenção especial, pois, embora tenha como objetivo a comunicação entre seres humanos, faz uso de diferentes instrumentos e canais, incluindo, simultaneamente, termos, definições, relações, símbolos, expressões lógicas, máquinas e programas.

A Lógica é a área que dá suporte à representação do conhecimento em ambientes virtuais, beneficiando diretamente a construção semântica das Ontologias (Guarino, 1997; Gruber, 2007; entre outros). Sendo a Lógica reconhecida atualmente como uma disciplina basicamente matemática, pode soar um pouco estranho ser ela incorporada como meio para tratar informação. É portanto lícito perguntar: como a lógica atua na ação representacional das Ontologias?

4.1 Lógica: área de abordagem da representação formal de enunciados

O termo “lógica” vem do grego clássico, λογική (logos), que significa palavra, pensamento, idéia. Sua utilização se deu de forma corrente a partir dos Estóicos, consagrando a Lógica como uma das três espécies de Filosofia. Supõe-se, pelos textos antigos, que o termo Lógica surgiu na época de

Aristóteles, podendo ter sido empregado pelos comentadores do filósofo como contraponto à Dialética encontrada na sua obra *Organon* (LALANDE, 1993).

Mesmo não havendo registros que comprovem sua participação na criação do termo, Aristóteles é considerado o precursor do pensamento lógico, sobretudo porque foi a partir de sua obra *Organon* que os métodos, conceitos, natureza e técnicas da Lógica foram discutidos pioneiramente, bem como a discussão sobre a Lógica Formal e a Material.

Encontram-se na literatura três períodos que marcaram a história da Lógica (TUGENGHAT, 1996). Um primeiro, que abrangeu a lógica antiga, da época de Aristóteles. Um segundo, denominado de Lógica moderna, em que os estudos se voltaram para as entidades dos conteúdos apresentados, caracterizados pela “predominância de problemas ligados à Teoria do conhecimento e à Psicologia, razão pela qual a pesquisa lógica, no sentido estrito, assim como o esclarecimento de conceitos lógicos básicos passaram a segundo plano”, enfraquecendo a área (TUGENGHAT, 1996, p.10). Embora tenha sido considerado um momento de retrocesso para a Lógica, foi nesse período que grandes filósofos se firmaram, a exemplo de Kant e Hegel, que fortaleceram as bases lógicas da área filosófica.

O último período começou com Frege, em 1879, se estendendo até os dias de hoje. Nele abandonou-se a Teoria do conhecimento e a Psicologia, retomando-se as bases lógicas de Aristóteles. A abordagem, por enfatizar a matemática e os símbolos e os cálculos formais, lhe rendeu o nome de Lógica matemática ou simbólica.

Considera-se que a expressão “Lógica formal” dá o sentido mais completo da Lógica. Ela é entendida como o

“estudo dos conceitos, juízos e raciocínios, considerados nas formas em que são enunciadas, e abstração feita da matéria à qual se aplicam, com vistas a determinar *in abstracto* as suas propriedades, a sua validade, os encadeamentos, e as condições sob as quais se aplicam ou se excluem uns aos outros.” (LALANDE, 1993, p. 630)

João Branquinho (2006) ressalta que a principal característica dessa disciplina é o fato de suas análises considerarem a forma do raciocínio, desconsiderando o conteúdo. Nessa medida, a Lógica formal estuda o conhecimento, do ponto de vista operatório, sem se deter sobre a validade do conteúdo pensado em si mesmo e ou sobre a sua entidade psíquica. É como dizer que $2 + 2$ é igual a 4, não importando a essência psíquica disso, e sim a forma e o que está escrito.

Pode-se dizer, portanto, que embora a Lógica esteja fortemente ligada à Filosofia ela é hoje uma disciplina de natureza matemática que concretiza, por meio de regras e esquemas formais, a inferência e o argumento, a manifestação e a validação do pensamento e do conhecimento humanos.

Em inferências de interesse lógico, é necessário dispor de sentenças ou premissas que venham de abstrações do conhecimento de uma certa comunidade de falantes.

Os diversos tipos de inferências vêm da oscilação entre Lógica formal e informal, cujas percepções de certo ou errado resultam de regras ou leis que orientam as inferências e os argumentos das premissas, regras essas que determinam estritamente se uma inferência é indutiva ou dedutiva.

A inferência indutiva se caracteriza nos casos em que a segunda premissa ou sentença sempre generaliza as informações contidas nas frases a partir das quais se fazem as inferências, como exemplificado abaixo (JOÃO BRANQUINHO, 2006):

- (a) Até agora, 2006, nenhuma amostra de água deixou de ferver quando aquecida a 100°C ;
- (b) A água ferve a 100°C .

Esse tipo de inferência é regido pela lógica indutiva e pela estatística, sendo válido probabilisticamente.

A inferência dedutiva se caracteriza pela forma lógica e particular em que as sentenças se apresentam para se inferir. Se faz exclusivamente através da Lógica dedutiva sem a noção de generalização que se encontra na indução. Seu raciocínio parte e conclui a partir das informações contidas na própria sentença inicial da inferência. As inferências desse tipo podem ser acompanhadas das seguintes expressões: não; se; então; e; ou; se, e somente se; todos; alguns. No entanto, o que marca sua essência é o compromisso com a verdade, como os exemplos abaixo.

- (a) Todo homem é mortal;
- (b) Sócrates é um homem
- (c) Então, Sócrates é mortal

- (a) Se neva, faz frio
- (b) Está nevando
- (c) Faz frio

As inferências dedutivas, como prezam a verdade dos fatos apresentados, validam ou não as premissas colocadas, dependendo da veracidade encontrada. A validação se dá por meio de regras de inferência diretamente associadas às expressões apresentadas acima, bem como pela lógica do raciocínio.

- (a) Nenhum dos amigos de Juan é meu amigo
- (b) Nenhum dos amigos de Juan é simpático
- (c) Portanto, nenhum dos meus amigos é simpático.

Neste exemplo, pode-se perceber a falta de lógica da conclusão. Não há, neste caso, conexão entre as premissas e a conclusão.

Allwood (1981, p.27) afirma que os aspectos mais importantes dos estudos da Lógica, referem-se ao “estúdio de las inferencias válidas y de los enunciados que son necesariamente verdaderos”. O autor afirma ainda que as inferências dedutivas são estudadas em maior profundidade do que as indutivas e trazem maiores contribuições para os estudos da linguagem, uma

vez que validam ou não as operações duvidosas da língua, as quais necessitam de argumentos lógicos para dissipá-las.

O argumento é justamente o raciocínio lógico que opera no ato de inferir. Da mesma forma que as inferências, o argumento também se divide em dois tipos: indutivo e dedutivo, seguindo as mesmas características dos respectivos tipos de inferências. Ou seja, o raciocínio indutivo opera por probabilidade, ao passo que o dedutivo, na maioria das vezes, justifica na íntegra todo o processo de inferência.

João Branquinho detalha, ao mesmo tempo, os constituintes de uma inferência.

Um argumento, dedutivo ou indutivo, é composto por um conjunto de frases a que chamamos premissas, por uma frase a que chamamos conclusão e por uma expressão que representa a relação que se reclama existir entre as premissas e a conclusão, p.ex., a expressão “logo” – a qual traduz a expressão latina “ergo”. (JOÃO BRANQUINHO, 2006, p.43)

Como se pode perceber, as expressões que determinam um argumento também estão presentes nas inferências dedutivas, marcando mais uma vez a valoração desse tipo de inferência para a lógica. São expressões que conectam e descrevem o tipo de relação existente entre as premissas e a conclusão.

Cabrera (2007, p.17), apresenta outras expressões para esse tipo de relação, afirmando que as inferências e, conseqüentemente os argumentos, são baseados em conexões lexicais com ênfase na linguagem comum. Em um de seus exemplos, considerou uma ação do dia a dia de dois amigos que falavam sobre uma certa mudança.

- (a) Não posso ir à tua casa hoje, estou muito ocupado.
- (b) Mas você disse ontem que me ajudaria a carregar meus móveis na minha mudança. Ontem, domingo, você não mencionou que teria outras coisas para fazer.

Nessa breve e simples troca os participantes fizeram uso de várias conexões, dando pé, pelo menos, às seguintes inferências lexicais:

- (1) A não mora junto com B; A está em sua própria casa hoje t1; portanto, A não está na casa de B hoje em t1.
- (2) A falou com B ontem; ontem foi domingo; portanto, A falou com B dois dias antes de uma terça-feira.
- (3) A está muito ocupado hoje; portanto, A tem muitas coisas a fazer que lhe tomarão a maior parte do seu tempo, ou todo ele.
- (4) B tem uma mudança para fazer hoje (dia s); portanto, B tem algo a fazer hoje (dia s).

Segundo o autor, como são ações que fazem parte do dia a dia de um falante da língua portuguesa, a intuição pode ajudar na validação e correção das premissas e conclusões através das conexões lexicais apresentadas nas inferências.

Os elementos necessários para a validação de um raciocínio, lexicais ou não, são constituintes do que Allwood (1981, p.28) chama de “Forma Lógica de los razonamientos”, que faz de qualquer ato ou conhecimento humanos um objeto formal passível de verificação e validação.

No caso das Ontologias, a validação do raciocínio por meio de elementos lógicos é de suma importância; deles dependem a qualidade das Ontologias, uma vez que se podem testar, por exemplo, conceitos contraditórios.

4.2 Formalismo: conceito, histórico e técnica

Para Salmon (2002), a possibilidade de expressar o pensamento, ato ou conhecimento humano por uma sentença de predicados, faz da Lógica um instrumento aplicável a diferentes áreas que necessitem tanto de um esforço intelectual para a apuração de provas, para distinguir entre o verdadeiro e o falso, como para tratar eficazmente o conhecimento.

A formalização apresenta uma funcionalidade que “designa a tendência a ocupar-se principalmente, se não exclusivamente, de caracteres formais”, sugerindo a ênfase na “forma do real frente a sua matéria ou frente a sua virtualidade” (MORA, 1971, p.719). Para Lalande (1993, p.427), o formalismo é uma “doutrina que consiste em afirmar que as verdades desta ou daquela ciência são puramente formais e repousam unicamente sobre convenções ou sobre definições de símbolos”. Ou seja, enfatiza-se prioritariamente o ponto de vista formal de uma determinada coisa, seja esta um conteúdo, um objeto ou uma disciplina.

Na literatura matemática, o termo formalismo é encontrado em três versões ou acepções. Na primeira e mais antiga versão, o formalismo se divide em duas vertentes: uma defende que as proposições da matemática são somente seqüências de símbolos, em que se fixam as regras de movimentação desses símbolos, sem a preocupação do sentido que formam a cada movimento, ou seja, em que a interpretação dessa dinâmica é irrelevante; a outra adota a doutrina de que “a existência de um objeto é garantida pela demonstração de consistência do sistema onde o objeto é representado” (JOÃO BRANQUINHO, 2006, p.359). Segundo o autor, o formalismo se associa ao critério de não-contradição.

Na segunda versão, o termo é usado para estabelecer um conjunto de doutrinas conhecidas por *Programa de Hilbert* (FERREIRA, 1995). O programa foi proposto por David Hilbert, em 1921, para reformular com rigor as bases da Matemática. Para ele, a Matemática poderia ser reduzida a um número finito de axiomas e, assim, qualquer proposição matemática poderia ser provada dentro de um sistema considerado completo. Mas esse programa foi contestado pelo matemático Kurt Gödel que, por meio do Teorema da incompletude (1931), provou ser impossível tal afirmativa. Gödel argumentou apresentando a tese de que um sistema axiomático consiste justamente em não poder provar sua própria consistência, do contrário ele só poderia ser inconsistente.

Na sua terceira e última versão, em meados dos anos 1970, a teoria godeliana ganhou forma e força dentro da expressão “doutrina formalista-positivista”, oriunda dos estudos de Georg Kreisel. Segundo este pensador, o

formalismo-positivista conseguiu se instalar na filosofia da matemática após sucessos parciais de formalização na área. O aspecto crucial dessa última concepção foi a rejeição de qualquer validade do conhecimento sobre conceitos abstratos, ou seja, de “extrapolações meramente verbais sobre o verdadeiro conhecimento de objetos e fatos concretos” (JOÃO BRANQUINHO, 2006, p.360). Dessa forma, o ganho do formalismo com tal doutrina, foi a substituição do emprego de conceitos abstratos, até então vigentes, para concepções que pudessem ser controladas por meio de um sistema formal, de maneira não intuitiva.

Para alguns matemáticos esta mudança não foi bem vista pois acreditaram que “a matemática formalista [reduziu-se] a conceitos para a compreensão dos quais é suficiente possuir uma lista de regras formais que os descrevessem integralmente” (JOÃO BRANQUINHO, 2006, p. 360)

No entanto, a doutrina formalista-positivista contribuiu para o fortalecimento da confiabilidade no conhecimento obtido por meio de regras formais, uma vez que os conceitos abstratos da matemática clássica sempre colocaram dificuldades em relação à compreensão e à confiabilidade da intuição abstrata. Criou-se, desse modo, a necessidade de controle de nossas concepções intuitivas por meio da formalização.

Embora tendo natureza matemática e filosófica, como exposto acima, a essência lógico-formal para expressar o conhecimento humano auxilia sobremaneira a área da linguística já há algum tempo. O Gerativismo é, de fato, uma espécie de formalismo que representou a volta à antiga tradição da Gramática Universal, que tem como pai Noam Chomsky, teórico que construiu as bases da Linguística estrutural e matemática contemporâneas (FARIAS JR., 2003; KENEDY, 2008).

Os gerativistas, em oposição aos behavioristas, acreditavam que o ser humano era capaz de construir e reconstruir constantemente frases e idéias novas, não dependendo unicamente do condicionamento social. Para eles, é um fenômeno nato ao ser humano. Com essa idéia, os gerativistas buscavam na linguística “construir um modelo teórico capaz de descrever e explicar a

natureza e o funcionamento dessa faculdade, o que significa procurar compreender um dos aspectos mais importantes da mente humana." (KENEDY, 2008, p. 129).

Com esse pensamento, teóricos gerativistas são motivados a estudar o funcionamento da linguagem na mente, através de investigações que analisavam a linguagem dentro de parâmetros gramaticais, matemáticos e abstratos, formalizando os estudos nessa área.

Tais análises abordavam a estrutura gramatical das frases, incluindo aspectos semânticos nas chamadas estruturas profundas e aspectos fonéticos nas estruturas de superfície. Em outras palavras, as interpretações semânticas ou significados de uma expressão são vistas como inerentes à estrutura profunda, e a construção da representação fonética dessa expressão, inerente à estrutura superficial (CHOMSKY, 1980, p.65-69). Diante dessa categorização frasal é possível formalizar os aspectos sintáticos e semânticos da linguagem.

Assim, Chomsky, em seu trabalho intitulado *Syntactic Structure* (1957), além de consolidar os estudos sobre os aspectos sintáticos da linguagem formal, mostra também a integração semântica da linguagem natural com as análises lógico-formais, ressaltando a importância de ver toda e qualquer linguagem como um objeto formal. O autor busca, com isso, na recursividade sintática a propriedade formal que dê conta da aparente infinitude de possibilidades de geração de frases e sentenças, proporcionando o que ele chama por "capacidade gerativa da linguagem" (CHOMSKY, 1957 apud FARIAS JR, 2003), a base que deu origem à gramática gerativa.

Já na década de 70, diante desses acontecimentos e estudos, viveu-se um momento em que a integração semântica da linguagem natural e as análises lógico-formais, inclusive de cunho filosófico, já estavam consolidadas - contribuições dadas por Chomsky, com os estudos sobre a linguagem natural e sua possível natureza formal -, influenciando a reestruturação da linguagem utilizada na lógica, a linguagem formal, que dá base também para a área de computação.

A linguagem formal, até a referida reestruturação, era sinônimo de “sistema sintático não-interpretado” (JOÃO BRANQUINHO, 2006), pois a preocupação era unicamente com as questões de sintaxe da linguagem, por ter sido este o primeiro problema reconhecido pela área de computação em relação às programações. Com a interferência de uma semântica lógica, essa linguagem passou a considerar também a interpretação como parte integrante de seu processo, auxiliando hoje na construção da nova WEB e seus aparatos tecnológicos, como as Ontologias.

4.3 Linguagens Formais para Construção de Ontologias: conceitos e técnicas

O principal objetivo dos sistemas lógicos formais é representar o conhecimento humano por meio de inferências e argumentos, representação essa que se utiliza de uma linguagem formal na qual os seus aspectos estruturais são representados por esquemas e categorias bem definidos.

Quem primeiro visualizou uma linguagem artificial ou formal foi o filósofo alemão G. W. von Leibniz, em meados do século XVII, com a intenção de criar uma língua filosófica que refletisse o pensamento humano. Mortari (2001,p.33) explica que sua proposta era o “desenvolvimento de um *calculus ratiocinator*, um calculo que permitiria tirar automaticamente conclusões a partir de premissas representadas na *lingua philosophica*”. Assim, bastaria tomar qualquer sentença de uma discussão filosófica, traduzir para a referida linguagem e calcular para se obter certa conclusão, como nos cálculos matemáticos. O propósito do filósofo era, de fato, sintetizar o pensamento humano em expressões matemáticas de fácil assimilação.

Mesmo tendo contribuído para o aparecimento de tal linguagem, Leibniz não a desenvolveu. Somente no século XIX a linguagem passou a ser um instrumento da Lógica, inicialmente com os trabalhos de G. Boole e, depois, com Gottlob Frege, com a publicação da obra *Conceitografia*, de 1879.

Em meados da década de 1950, foi criada a teoria das Linguagens Formais. Menezes (2008) afirma que a princípio ela foi pensada apenas para se trabalhar com linguagens naturais, no entanto, logo se percebeu sua utilidade para operacionalizar as linguagens artificiais voltadas para a computação.

Desde então, o estudo das linguagens formais desenvolveu-se significativamente, com destaque para aplicações em *análise léxica* e *análise sintática* de linguagens de programação, modelagem de *circuitos lógicos* ou *redes lógicas*, de *sistemas biológicos*, entre outros enfoques. Mais recentemente, destacam-se aplicações relacionadas com *sistemas de animação*, *hipertextos* e *hipermídias*, bem como o tratamento de *linguagens não-lineares*, como *linguagens planares*, *linguagens espaciais* e *linguagens n-dimensionais*. (MENEZES, 2008, p. VII)

Esses investimentos foram tão importantes que atualmente é impensável uma lógica sem linguagem formal.

As linguagens formais são entendidas aqui como:

linguagens artificiais construídas pelos lógicos com o objetivo, científico, de estudar conceitos lógicos fundamentais (p. ex. verdade, validade ou consequência, consistência, completude, correção, decidibilidade) e com o objetivo, digamos, pedagógicos, de expor a teoria lógica (JOÃO BRANQUINHO, 2006, p.442).

Por ser uma linguagem que expõe o raciocínio lógico, pode-se dizer que se trata de um elemento que materializa a abstração da lógica em símbolos e expressões legíveis por um falante da língua natural.

Isso posto, infere-se que a linguagem é o elemento mais essencial na área da computação, sobretudo a formal ou artificial, pois dela depende a comunicação do homem com a máquina. Para isso, trabalha diretamente com a sintaxe e a semântica. Para melhor compreensão do que foi afirmado, se faz necessário uma breve introdução aos conceitos de linguagem.

A linguagem, em sentido amplo, é entendida como um instrumento auxiliar na função de “expressar o sistema lingüístico fixado numa determinada sociedade” (LALANDE, 1993), utilizado para a comunicação dos falantes de uma mesma língua.

Para Mortari (2001), a linguagem é constituída por um “conjunto finito de elementos básicos” com os quais se formam inúmeras expressões linguísticas que constroem as palavras e as frases. Esses elementos básicos correspondem aos morfemas e à morfologia, ou seja, às menores unidades de significação e às regras gramaticais de construção de expressões linguísticas.

Combinações de morfemas, de acordo com certas regras (a morfologia), nos permitem formar palavras, e combinações de palavras, de acordo com certas outras regras (a gramática), nos permite formar frases e sentenças. Com as sentenças, claro, você pode construir estruturas mais complexas, como argumentos, discursos, diálogos, artigos de jornal, etc. (MORTARI, 2001, P.32)

Mortari (2001) ainda aponta os três níveis dos estudos da linguagem: a sintaxe, a semântica e a pragmática.

O primeiro nível, a sintaxe, diz respeito à formação das palavras e das frases, ou seja, às regras gramaticais que regem a composição dessas unidades da língua. Nesse nível não se estudam os significados constantes desses construtos, porém apenas a parte formal dos símbolos. Os estudos de sintaxe foram os primeiros a contribuir com a área de computação, pelos motivos que vimos acima. Como consequência, os estudos dessa área tiveram

um desenvolvimento considerável, dotados, portanto, de construções matemáticas bem definidas e universalmente conhecidas.

O segundo nível, a semântica, é responsável justamente pelos estudos dos significados de palavras e frases. Nele procura-se entender o sentido de cada símbolo da forma lingüística, por meio da interpretação, sem se preocupar com a interação entre esses significados e o contexto social em que são usados. As abordagens semânticas são recentes na área de computação; datam, do final da década de 1990, quando tiveram início os estudos preliminares para a Web semântica.

O ultimo nível apontado pelo autor, a pragmática, é responsável pelos estudos de contextos culturais, aspecto não considerado na semântica. Ou seja, estuda-se a formação das frases pelos falantes de uma determinada língua e cultura, sendo esses os indicadores do sentido social da frase construída. Pensando de forma pragmática, a realidade de cada contexto social e cultural é responsável por essa construção, partindo dos elementos dinâmicos que existem dentro das relações humanas e sociais, incluindo a linguagem como instrumento para tais relações e as palavras e frases como insumos da linguagem.

Como se pode perceber, a semântica e a pragmática, embora bastante próximas, são elementos lingüísticos bastante diferentes, uma vez que o primeiro se preocupa exclusivamente com o significado real da sentença e, o segundo, com um sentido secundário em que as frases expressam de acordo com a realidade de cada falante. Mortari (2001) exemplifica esta afirmativa com a seguinte frase: “Você sabe que horas são?”. No sentido semântico pode-se concluir que o falante da sentença está perguntando apenas se o dono do relógio sabe que horas são, tendo assim como possível resposta um “sim” ou um “não”. Contudo, no sentido pragmático, espera-se um pouco mais do que se perguntou, algo como as horas completas do momento em que a pergunta foi realizada.

Para Menezes (2008), na linguagem formal, a preocupação se direciona basicamente para a sintaxe e a semântica, exigindo regras para compor

símbolos com maior exatidão (ou, pelo menos, tentam), diferentes das sublimações da língua natural dos falantes. Ou seja, procura-se elaborar regras gramaticais rigorosamente definidas para diminuir as várias nuances de uma mesma interpretação.

A nosso ver, a pragmática é importante para a representação de conteúdos em sistemas, inclusive em Ontologias, onde a precisão da recuperação das informações se deve ao contexto social e cultural dos que solicitam informações. Mesmo sendo a pragmática linguística um elemento importante para as linguagens formais, elas não evoluem de acordo com cada grupo que a utiliza, como ocorre na linguagem natural. Ela é estabilizada, o que facilita seu manuseio e entendimento para a padronização do tratamento de informação. Sua formação é proveniente de um alfabeto, formado por caracteres que não correspondem a letras, mas a símbolos, e por uma gramática, que são regras de construção de expressões formais, como os termos e as fórmulas, típicos da área de computação (MENEZES, 2008).

As linguagens formais são tratadas e estudadas de acordo com uma hierarquia estabelecida por Noam Chomsky, por isso conhecida por Hierarquia de Chomsky (MENEZES, 2008), que estabeleceu as classes como possíveis modelos para a linguagem natural. São elas:

a) Linguagens Regulares (LR) ou Tipo 3.

Menezes (2008, p.41) explica que esse tipo de linguagem é a mais simples existente na área computacional, embora tenha a capacidade de desenvolver algoritmos que reconhecem ou geram objetos digitais, sejam esses de que natureza for, ou seja, imagem, texto, som, entre outros; bem como, algoritmos de conversão entre formalismos de pouca complexidade e de fácil implementação. Entretanto, por serem linguagens simples, apresentam limitações para uso em projetos computacionais mais elaborados. Por exemplo, a maioria das linguagens de programação de propósitos gerais como Pascal, C, Java, entre outras de nível mais elaborado, não podem ser regulares.

b) Linguagens Livres do Contexto ou Tipo 2.

Essas linguagens são de importantes para a computação, pois se constituem num universo maior de linguagens do que as regulares, incluindo as de programação como a Pascal, C, Java, entre outras. Têm essa denominação por serem geradas por alguma gramática livre de contexto. Seus algoritmos de reconhecimento e geração são bastante simples porém eficientes. Suas aplicações e resultados são direcionados para as linguagens artificiais (formais), em especial as linguagens de programação. No entanto, é uma classe restrita de linguagens, sendo de fácil definição as que não pertencem a ela.

c) Linguagens Sensíveis ao Contexto ou Tipo 1.

São linguagens geradas por alguma gramática sensível ao contexto, bem como aceita por um autômato ou dispositivo linearmente limitado, ou seja, são aquelas que podem ser aceitas por uma Máquina de Turin com Fita limitada. Menezes (2008, p.171) define Máquina de Turin como “uma formalização de um procedimento efetivo (algoritmo ou função computável), ou seja, uma seqüência finita de instruções, as quais podem ser realizadas mecanicamente, em um tempo finito”. No que diz respeito à Linguagens formais, é uma classe importante na Hierarquia de Chomsky, aplicadas na maioria das linguagens de computação e na lógica.

d) Linguagens Recursivamente Enumeráveis ou Tipo 0

As linguagens formais são denominadas recursivamente enumeráveis quando são computáveis, ou seja, quando há um algoritmo que faz o reconhecimento dessa linguagem. Uma linguagem só é computável quando é aceita por uma Máquina de Turing, que controla as paradas para as instâncias positivas de um projeto computacional, com um "sim" ou o número "1". Menezes (2008) complementa essa afirmação dizendo que essa máquina se constitui no mais genérico dispositivo computacional e, no caso dessa linguagem em especial, um autômato

que possui uma fita sem um comprimento definido como tamanho máximo, em outras palavras, uma fita infinita. Desta forma, pode-se dizer que as linguagens recursivas são todas as linguagens reconhecidas mecanicamente, incluindo todas as mencionadas acima.

Allwood (1981, p.9) elenca vários símbolos utilizados nessas linguagens formais, com suas respectivas notações. Eis alguns exemplos de como esses símbolos são trabalhados numa linguagem formal (Menezes, 2008, p.34).

a) O conjunto vazio $\emptyset$ e o conjunto formado pela palavra vazia $\{\epsilon\}$ são linguagens sobre qualquer alfabeto. Obviamente, vale que: $\emptyset \neq \{\epsilon\}$

b) Os conjuntos Σ^* e Σ^+ são linguagens sobre o alfabeto Σ qualquer. Obviamente, vale que: $\Sigma^* \neq \Sigma^+$

c) Suponha o alfabeto $\Sigma = \{a,b\}$. Então, o conjunto de palíndromos (palavras que tem a mesma leitura da esquerda para a direita e vice-versa) sobre Σ é um exemplo de linguagem infinita. Assim, são palavras dessa linguagem: ϵ , a, b, aa, bb, aaa, aba, bab, bbb, aaaa,...

d) Uma linguagem formal L sobre um alfabeto Σ , é um conjunto de palavras sobre Σ , ou seja: $L \subseteq \Sigma^*$

O interesse para o uso dessa linguagem formal é que o raciocínio originalmente apresentado em uma língua natural é transcrito ou traduzido para esses símbolos, totalmente especificados e fixados, que exprimem um entendimento alcançável, em princípio, por qualquer falante de qualquer idioma, evitando assim as ambiguidades próprias da língua natural.

Existem vários sistemas lógicos que se utilizam dessa linguagem formal. No caso das Ontologias, é utilizada para a codificação do processo de representação formal do conhecimento. São linguagens utilizadas em operadores, axiomas, declarações, entre outros. Almeida e Bax (2003) elencam algumas das linguagens utilizadas na representação formal do conhecimento em Ontologias:

Cycl – Linguagem formal que expressa conhecimento por meio de um vocabulário de termos (constantes semânticas, variáveis, números, sequências de caracteres etc.) os quais são combinados em expressões, sentenças e finalmente bases de conhecimentos (LENAT & GUHA, 1990)

Flogic (Frame Logic) – Integra frames e lógica de primeira ordem. Trata de forma declarativa os aspectos estruturais das linguagens baseadas em frames e orientadas a objetos (identificação de objetos, heranças, tipos polimórficos, métodos de consulta, encapsulamento, etc). Permite a representação de conceitos, taxonomias, relações binárias, funções, instâncias, axiomas e regras (KIFER, LAUSEN & WU, 1990)

LOOM – Descendente da família do KL-ONE (Knowledge Language One), é baseada em lógica descritiva e regras de produção. Permite a representação de conceitos, taxonomia, relações n-árias, funções, axiomas e regras de produção (BRILL, 1993)

CARIN – Combinação da Datalog (Linguagem baseada em regras) e lógica descritiva ALN. Uma Ontologia CARIN é construída por dois componentes terminológicos: um conjunto de conceitos com declarações de inclusão e um conjunto de regras que usam os conceitos (LEVY & ROUSSET, 1996)

GRAIL – Linguagem que especifica uma Ontologia do domínio médico (Galen). Baseada em lógica descritiva, terminologicamente limitada, permite a construção de hierarquias de primitivas e axiomas de inclusão de conceitos (RECTOR et alii, 1997).

ONTOLÍNGUA – Combina paradigmas das linguagens baseadas em frames e lógica de primeira ordem. Permite a representação de conceitos, taxonomia de conceitos, relações n-árias, funções, axiomas, instâncias e procedimentos. Sua alta expressividade causa problemas na construção de mecanismos de inferências (CHAUDHRI et alii, 1998).

OCML – Permite a especificação de funções, relações de classes, instâncias e regras. Utilizadas em aplicações de gerenciamento do conhecimento, desenvolvimento de Ontologias, comércio eletrônico e sistemas baseados em

conhecimento. Aplicada em Medicina, Ciências Sociais, Memória Corporativa, Engenharia, portais da Web etc. (DOMINGUE, MOTTA & CORCHO, 1999).

OML (Ontology Markup Language) – Linguagem baseada em lógica descritiva e grafos conceituais que permite a representação de conceitos organizados em taxiomas, relações e axiomas (KENT, 1999).

RDF (Resource Description Framework) / **RDFS** (RDF Schemas) – Desenvolvido pela W3C, tem por objetivo a representação de conhecimento por meio de redes semânticas. Permitem representar conceitos, taxonomia de conceitos e relações binárias (LASSILA & SWICK, 1999).

NKRL (Narrative Knowledge Representation Language) – Linguagem de representação baseada em frames, especialmente desenvolvida para descrever modelos semânticos de documentos multimídias (BERTINO, BARBARA & ZARRI, 1999).

SHOE (Simple HTML Ontology Extensions) – Utiliza extensões ao HTML, adicionando marcações para inserir metadados em paginas Web. As marcações podem ser utilizadas para a construção de Ontologias e para anotações em documentos da Web. (HEFLIN & HENDLER, 2000)

XOL – Linguagem que pode especificar conceitos, taxonomias e relações binárias. Não possui mecanismos de inferência, tendo sido projetada para o intercâmbio de Ontologias do domínio da biomedicina (KARP, 1997)

OIL (Ontology Interchange Language) – Precursor do DAML+OIL e base para uma linguagem para a Web semântica. Combina primitivas de modelagem das linguagens baseadas em frames com a semântica formal e serviços de inferências da lógica descritiva. Pode verificar classificações e taxonomias de conceitos (FENSEL et alii, 2001)

DAML (DARPA Agent Markup Language) + **OIL** – Linguagem de marcação semântica para a Web que apresenta extensões a linguagens como a DAML,

RDF e RDFS, por meio de primitivas de modelagens baseadas em linguagens lógicas (HORROCKS et alii, 2001).

FOML (Formal ontology markup language) –Linguagem de marcação, baseada em XML, que conecta documentos da Web com Ontologias formais. O objetivo é a aquisição automática de conhecimento de domínios específicos. (OGATA, 2001).

KIF (*Knowledge Interchange Format*): Linguagem lógica de primeira ordem anotônica, com sintaxe simples e algumas extensões para suportar um raciocinador de relações. (CANTELE et alii, 2010).

As linguagens de representação de conhecimento acima têm em comum: a delimitação de vocabulários de termos, tratados de forma declarativa, e regras para especificar relações entre os termos. São baseadas em semânticas formais que permitem realizar inferências.

5. ANÁLISE SEMANTICA DA ONTOLOGIA RADLEX

5.1 Percurso metodológico

Essa pesquisa se situa no âmbito das Ciências Sociais e busca responder aos questionamentos direcionados aos instrumentos de organização e tratamento da informação em ambientes virtuais, mais precisamente as ontologias. Caracteriza-se como qualitativa, que segundo Maanen (1979 apud NEVES, 1996, p.1) “compreende um conjunto de diferentes técnicas interpretativas que visam descrever e analisar os componentes de um sistema complexo de significados”, cuja finalidade busca “expressar o sentido dos fenômenos do mundo real, (...) reduzindo a distância entre (...) teoria e dados”.

Tomando por base a teoria peirceana, procurou-se identificar como ocorrem os processos de significação das ontologias, tendo como categoria principal o interpretante peirceano, elemento pragmático da teoria, responsável pela produção de sentidos dos objetos representados. Foram igualmente identificadas as operações lógicas de construção de Ontologias de domínio (arquitetura e elementos de composição). O percurso pode ser sintetizados nos seguintes passos:

- Compreensão dos principais processos de significação de uma Ontologia de domínio, verificando a sintaxe e a semântica das relações existentes entre termos.
- Aplicação da teoria do interpretante aos processos de significação das Ontologias.

- Sistematização dos parâmetros teóricos e metodológicos de construção de Ontologias.

Para tal, realizou-se a pesquisa direta em uma ontologia de domínio selecionada para o estudo, por meio de observação intensiva para coletar dados. Segundo Lakatos; Marconi (2001, p.190) esta “técnica de coleta de dados para conseguir informações utiliza os sentidos na obtenção de determinados aspectos da realidade. Não consiste apenas em ver e ouvir, mas também em examinar fatos ou fenômenos que se desejam estudar”. Através dessa técnica foi realizada a delimitação dos termos coletados para fins de análise, uma vez que a ontologia utilizada como estudo de caso é de grande porte, contendo um número bastante significativo de termos e relações.

5.2 Universo e amostra

Para a delimitação do universo foi feita primeiramente a escolha da ferramenta de construção de ontologias, para então selecionar a ontologia da pesquisa. Assim, foi realizada uma pesquisa primeiramente sobre os tipos de ferramentas de construção de ontologias existentes, através da qual foi selecionada a ferramenta Protégé 2000, por ser de acesso gratuito, ter uma interface de fácil manuseio e por ser uma das mais utilizadas para a construção de ontologias nos dias atuais.

Dentre os tipos de ontologias que utilizam a plataforma Protégé, foram escolhidas as ontologias de domínios, por representarem domínios de conhecimento que necessitam de uma representação bem definida de assuntos. Diante dessa escolha, foi selecionada uma ontologia específica da área de radiologia, Radlex Ontology, por integrar um número considerável de terminologias radiológicas em seu vocabulário, por ser uma ontologia complexa, de grande porte e em constante desenvolvimento e avaliação.

Logo após definida a ontologia para análise, procedeu-se ao processo de observação de suas categorias e respectivas classes, para definir a seleção das classes para coleta. Assim, foram selecionadas as quatro classes da ontologia que a nosso ver são representativas de sua composição e complexidade:

➤ Class Radlex Descritor

➤ Subclasses selecionadas do Radlex Descritor:

- Image Observation Characteristic
- Pathophysiologic Process Modifier

➤ Class Radlex Entity

➤ Subclasses selecionadas do Radlex Entity:

- Anatomical Cluster
- Object

Foi em seguida, definida uma amostra final de 615 termos para análise.

5.3 Critérios de análise das relações semânticas

Para criar critérios de análise das relações semânticas foi adotada a norma ANSI.ISO Z39.19-2005, que propõe diretrizes de elaboração de vocabulários controlados em geral, inclusive ontologias. De acordo com as diretrizes propostas por essa norma, há três tipos básicos de relações entre termos:

↳ **relações hierárquicas**: baseadas em níveis de superordenação e subordinação, onde o termo superordenado representa uma classe ou um todo, e os termos subordinados referem-se a seus membros ou peças.

As Relações hierárquicas cobrem três situações logicamente diferentes e mutuamente exclusivas (ISO Z39.19, p.47-49, 2005. *tradução nossa*):

- a) **genérica** – aqui elas se dão a partir do princípio lógico de relacionamentos genéricos, que expressa a ligação entre uma classe e seus membros ou espécies. Esse tipo de relação utiliza frequentemente a expressão lógica “Is_a”. Ocorre sempre dentro de uma mesma categoria.
- b) **instância** – este relacionamento identifica a ligação entre uma categoria geral de coisas ou eventos, expressa por um substantivo comum, e uma instância individual da categoria, muitas vezes, um nome próprio. Este tipo de relacionamento também é representado por “Is_a”.
- c) **todo-parte** – abrange as situações em que um conceito é inerentemente incluído no outro, independentemente de contexto, de modo que os termos podem ser organizados em hierarquia lógica, com o todo tratado como um termo mais amplo.
- d) **polihierárquica** – alguns conceitos pertencem, por motivos lógicos, a mais de uma categoria. Nestes casos as relações são ditas polihierárquicas.

↳ **relações de equivalência**: quando o mesmo conceito pode ser expresso por dois ou mais termos, um deles é selecionado como preferencial e o outro como alternativo ou não preferencial. Essa relação entre os dois termos é uma equivalência, onde cada termo é considerado como uma referência ao mesmo conceito. A norma especifica ainda as seguintes formas encontradas para especificar essa relação: sinônimos, variantes lexicais, quase-sinônimos e postagem genérica (ISO Z39.19, p.43-45, 2005. *tradução nossa*):

- a) **Sinônimos**: são termos cujos significados são considerados como o mesmo ou quase o mesmo em uma ampla gama de contextos.
- b) **Variante lexicais**: diferem dos sinônimos, considerando que sinônimos são formas diferentes para o mesmo conceito, enquanto variantes lexicais são diferentes formas de palavras para a mesma expressão. Estas formas podem derivar de variações ortográficas ou gramaticais ou de formatos abreviados.

- c) **Quase-sinônimos:** são termos cujos significados são geralmente considerados como diferentes, mas que são tratados como equivalentes para efeitos de um vocabulário controlado. A caracterização como quase-sinônimos depende do domínio abrangido pelo vocabulário controlado e seu tamanho. Quase sinônimos podem incluir antônimos.
- d) **Postagem genérica:** técnica em que o nome de uma classe e os nomes dos seus membros são tratados como equivalentes, com o nome amplo da classe funcionando como o termo preferido.

↳ **relações associativas:** abrangem as associações entre os termos que não são nem equivalentes nem hierárquicos. São, no entanto, termos semântica ou conceitualmente associados a tal ponto que a ligação entre eles deve ser explicitada no vocabulário controlado, com o fundamento de que ele pode sugerir termos adicionais para uso na indexação ou recuperação. Na documentação tradicional são os chamados Termos relacionados.

Num segundo momento, foram relacionados os tipos que geralmente são utilizados nas ontologias. Nesta fase, foram selecionados os trabalhos de autores que permitissem sistematizar de forma coerente tais relações: Sales (2006), Smith (2005) e Sowa (2000).

A literatura mostra que nos ambientes virtuais não basta estabelecer relações semânticas entre objetos virtuais, é necessário descrever como essas relações são efetivadas, sendo de suma importância expressar literalmente a forma como essas ligações se dão isto é, “explicitar a própria relação” (SALES, 2006). Essa forma de apresentação é o que a autora denomina de expressão formal.

Dessa forma, conforme Sales (2006), “as relações nas ontologias se apresentam de duas formas: na terminologia - sistemas de termos com definições em linguagem natural; e na parte formal – expressas através de linguagem própria para entendimento da máquina”. A autora estipulou tais relações como categoriais e formais, definindo-as da seguinte forma:

- **relações categoriais**: são aquelas que revelam duplas categoriais. Ex.: coisa-processo, material-produto e outras duplas que categorizam as relações tradicionais de Vocabulários.

- **relações formais**: são aquelas que revelam o tipo de relações existentes entre as duplas de categorias. Ex.: causado por, ocorre, etc...

Nessa dualidade foram identificados três tipos de relações e suas respectivas categorias formais, totalizando seis tipos de relações: relações categoriais partitivas, relações formais partitivas, relações categoriais genéricas, relações formais genéricas, relações categoriais funcionais e relações formais funcionais.

Smith (2005, p.3) enfoca as relações existentes para uso exclusivo de ontologias, ou seja, “which we take to mean relations that obtain between entities in reality, independently of our ways of gaining knowledge about such entities”.

São listadas pelo autor as relações de uso geral nas ontologias da área biomédica, ou seja, aquelas que podem ser usadas em qualquer vocabulário biológico, em vez das relações específicas, tais como *genome_of* ou *sequence_of*, que se aplicam apenas às entidades biológicas de certos tipos. Assim, ele aponta três relações binárias para a construção semântica de um vocabulário ontológico:

- a) Class, class: a relação **is_a**:
- b) Instance, classe: **instance_of**
- c) Instance, instance: **part_of**

A partir dessas três especificações, Smith (2005) elabora uma tabela com as subdivisões das relações existentes nas ontologias biológicas. Para o autor, “Os itens dessa lista foram selecionados porque eles têm um alto grau de inteligibilidade para os autores humanos e curadores de ontologias biológicas. Para fins de aplicações de computador de apoio”. (SMITH, 2005. *Tradução nossa*)

5.4 Apresentação da Ferramenta PROTÉGÉ-2000

O Protégé foi desenvolvido pelo setor de informática médica da Universidade de Stanford, no final da década de 1980, para dar suporte a uma base de conhecimento especializado da área médica. Foi considerado originalmente como um projeto limitado, um sistema de banco de dados clássico com uma ferramenta de aquisição de conhecimento voltado apenas para a área específica a que servia, a oncologia (SEMPREBOM, et ali, 2007).

A ferramenta foi se modernizando, equiparando-se tecnologicamente aos sistemas baseados em conhecimento (SBC). Integrou diversos formalismos e estratégias de inferências, diversas tarefas num mesmo ambiente, permitindo a criação automática de formulários para entrada de conhecimento e acesso a Ontologias via Open Knowledge Base Connectivity (OKBC) (SEMPREBOM, et ali, 2007).

Para isso, mudou da linguagem Java, para um sistema de código aberto (*open source*), de modo a permitir seu uso por qualquer usuário e/ou grupo de pesquisa, fato que promoveu um desenvolvimento tecnológico significativo do programa.

Hoje, com sistema multiplataforma e adaptação para editar modelos em diferentes linguagens da *Web*, se comporta como um ambiente interativo de desenvolvimento de projetos de construção de Ontologias em diversas áreas do conhecimento. A premissa principal do atual Protégé é que os sistemas baseados em conhecimento são geralmente muito caros para serem construídos e mantidos (STANFORD, 2010). A expectativa é de a base do sistema ser desenvolvida pelo esforço de uma equipe diversificada e gratuita, incluindo programadores e especialistas internacionais. Apresenta, assim, uma arquitetura com interface de fácil manuseio para que usuários sem grandes conhecimentos de computação possam criar suas Ontologias e reutilizá-las para diminuir o tempo empregado no desenvolvimento e manutenção do programa. Ainda de acordo com os pesquisadores da Stanford, vários

aplicativos do programa podem usar o mesmo domínio da Ontologia para resolver problemas diferentes, podendo-se usar o mesmo método para a solução de problemas de diferentes Ontologias.

A ferramenta é basicamente organizada em módulos para que o programa possa se desenvolver livremente. Se existir algum problema em um determinado aspecto da estrutura, a verificação é feita somente nesse módulo, deixando os outros módulos em uso livre (SEMPREBOM et al, 2007).

Segundo os autores, a modularização é dividida em três partes: interface com o usuário, núcleo do Protégé, armazenamento persistente, conforme figura abaixo:

Figura 5: Módulos da Ferramenta Protégé 2000

Fonte: Semprebom, et al, 2007.

De acordo com site oficial do Protégé (STANFORD, 2010), o programa suporta duas maneiras principais de modelar as Ontologias:

- através do editor do Protégé-Frames
- através do editor Protégé-OWL

No editor do Protégé-Frames, um projeto de Ontologia está pautado num conjunto de classes organizadas dentro de uma hierarquia que representa um domínio de conceitos, bem como um conjunto de *slots* associados a essas classes para descrever as suas propriedades e relacionamentos e um conjunto de instâncias das classes, que contêm os valores específicos de cada propriedades estabelecida.

Em relação ao editor do Protégé-OWL, ele é visto como uma extensão do para desenvolver o aspecto semântico da web, em particular na W3C Web Ontology Language (OWL). Uma Ontologia desenvolvida em OWL pode incluir descrições de classes, propriedades e as suas instâncias, em que a semântica especifica como obter as consequências lógicas, ou seja, fatos literalmente não presentes, mas envolvidos pela semântica.

O modelo de Ontologias no Protégé é definido por meio de uma interface gráfica de fácil manuseio (conforme figura 6). Através dela se podem criar conceitos do domínio de interesse, sob uma estrutura que lembra o esquema dos instrumentos documentários (sistemas de classificação e tesauros): os conceitos são organizados em hierarquias de classes e subclasses, organizados em estrutura de árvore.

Figura 6: Interface da Ferramenta Protégé 2000

Fonte: http://protege.stanford.edu/doc/users_guide

A interface gráfica é baseada no padrão OKBC (*Open Knowledge-Base Connectivity*), um protocolo criado e bastante utilizado pela Universidade de Stanford. É uma interface de aplicação, conhecida na área da computação por API - *application-programming interface*, que possibilita o acesso à bases de conhecimento armazenadas em KRS - sistemas de representação de conhecimento. Em relação ao protocolo, os pesquisadores do laboratório de informática da Stanford explicam que:

O OKBC fornece um modelo uniforme de KRSs baseado em uma concepção comum de classes, indivíduos, entalhes, facetas, e heranças. OKBC é definido em uma linguagem de programação de forma independente, e tem implementações existentes em Common Lisp, Java e C. O protocolo suporta transparentemente em rede, bem como acesso direto à KRSs e bases de conhecimento. O OKBC consiste de um conjunto de operações que oferecem uma interface genérica para KRSs subjacente. Esta interface (...) permite o desenvolvimento de

ferramentas (por exemplo, os navegadores gráficos, editores de quadros, ferramentas de análise, ferramentas de inferência) que operam em muitos KRSs. Ela tem sido usada com sucesso em vários projetos em curso para SRI na Universidade de Stanford. (STANFORD, 2010)

Através dela é possível ter acesso a todas as partes do programa, sobrepondo guias para uma apresentação compacta e conveniente das peças para a co-edição entre elas. É um design de "abas" que integra os seguintes elementos:

- (1) **Ontologia** – define o conjunto de conceitos e seus relacionamentos
- (2) **ferramenta de aquisição de conhecimento** – projetada para ser de domínio específico, permitindo que os especialistas do domínio de forma fácil e naturalmente entre os seus conhecimentos da área.
- (3) **base de conhecimento** – usada com um método de resolução de problemas para responder perguntas e resolver problemas relacionados com o domínio.
- (4) **aplicações** – produto final criado quando a base de conhecimento é usada para resolver um problema do usuário final, utilizando adequado-problema de resolução de sistema especialista, ou de apoio à decisão métodos.

Para ter sucesso nos projetos implementados para a construção de um sistema de informação com o Protégé 2000, o trabalho deve ser visto como uma atividade artesanal, uma vez que o caráter altamente interativo e flexível dá liberdade de planejamento e execução para os desenvolvedores de Ontologias. Assim, é recomendada a revisão contínua durante todo o processo de construção da Ontologia para que a flexibilidade e a interatividade não sejam travas para o bom desempenho do projeto (STANFORD, 2010).

A imagem abaixo mostra o fluxo de desenvolvimento de um projeto Protégé, incluindo os ciclos de revisões necessários para o bom andamento do processo de construção. As setas pretas indicam o avanço do processo e as setas azuis os lugares exatos, cujas revisões se fazem necessárias.

Figura 7: Fluxo de desenvolvimento de um projeto Protégé

Fonte: http://protege.stanford.edu/doc/users_guide

5.5 Apresentação do Radlex Ontology

5.5.1 Conceito e Histórico

A padronização de linguagens documentárias com fins de recuperação da informação no meio digital já é realidade em várias áreas científicas, fato que impulsiona a criação de Terminologias e Ontologias.

A área médica é, atualmente, uma das mais avançadas em sistemas automatizados para armazenagem de informações hospitalares. Esse fato é proveniente da necessidade de informações atualizadas e organizadas sobre procedimentos, diagnósticos, relatórios hospitalares, bem como assuntos afins para tomadas de decisões emergenciais acerca da vida humana.

Na área de Radiologia, essa necessidade não é diferente, dado o aumento constante do volume de imagens produzidas em instalações clínicas. Por isso, é uma das áreas que está à frente no desenvolvimento de vocabulários terminológicos para o tratamento de imagens produzidas para laudos radiológicos. Essa tarefa inclui também a transposição de sistemas terminológicos tradicionais para sistemas *on line* com a finalidade de promover a comunicação em ampla escala. Arquivos de acesso local sem a interligação com outros vocabulários, não satisfazem mais a demanda da área.

Sistrom (2006 apud SANTOS E FUJINO, 2010) corrobora essa afirmativa ao explicar que “Nenhuma outra área de conhecimento da Medicina tem experimentado o aumento das facilidades disponíveis, em termos de volume e variedade, como é o caso da Radiologia, nas últimas décadas”

Esse histórico de iniciativas no âmbito do tratamento da informação pela Radiologia começou com o índice para diagnóstico radiológico, do *American College of Radiology Index of Radiological Diagnosis*, conhecido como ACR, um instrumento terminológico que serviu durante muitos anos para o tratamento da informação na área. Foi desenvolvido para categorizar e organizar assuntos relativos a imagens radiológicas (LANGLOTZ, 2006, p.1), e se mostrou eficiente, durante anos, por oferecer identificadores anatômicos e patológicos em códigos de números decimais, um método eficiente para a organização e recuperação de conteúdos em ambientes tradicionais. Com ele a recuperação de conteúdos era facilitada, pois continha uma organização que denotava o termo mais específico e o mais geral na taxonomia, oferecia um esquema de numeração no índice que garantia a comunicação dos materiais de ensino entre várias instituições, sendo, ainda, um esquema padrão de fácil entendimento em um ambiente convencional de informação.

Com o progresso tecnológico, gerando a transformação de materiais de imagem impressos para arquivos *online*, com ampla disseminação e produção no ambiente virtual, o Índice se tornou insuficiente para a recuperação de conteúdos da área. Assim, duas grandes limitações fizeram do índice ACR um instrumento obsoleto para a demanda científica radiológica atual: a) diferente dos computadores, o índice não tinha a capacidade de gerenciar mega listas

de termos e suas respectivas relações semânticas; b) não tinha a flexibilidade necessária para atualização diária de termos novos na lista, devido ao seu esquema fixo de numeração, que, projetado para o processamento de informações limitadas do ser humano, não proporciona a facilidade de mudança dos códigos pré-estabelecidos (LANGLOTZ, 2006).

Tais desvantagens levaram a comunidade radiológica a reivindicar um novo sistema terminológico, com viés tecnológico, atualizado, com cobertura mais abrangente de conteúdos, envolvendo os vários vocabulários existentes na área. Dessa forma, a Sociedade Radiológica da America do Norte – RSNA, criou o projeto Radlex para viabilizar a comunicação de conteúdos em radiologia, em grande escala, com a organização e a tecnologia necessárias para tratamento e recuperação eficazes.

O projeto Radlex Ontology é, por definição, um vocabulário terminológico especializado em Radiologia, entendido pelos especialistas como um instrumento que agrega valores conceituais da área radiológica, funcionando no fornecimento de uma “estrutura uniforme para capturar, indexar, e recuperar uma variedade de fontes de informação neste domínio, tais como arquivos de ensino, pesquisa de dados, e relatórios de exames” (CASTILLAS, 2007).

O primeiro projeto Radlex foi desenvolvido em processadores de texto e planilhas, ou seja, em Word e Excel, um formato que logo se mostrou insuficiente para acesso e gerenciamento dos conteúdos, devido alguns entraves evidentes:

RadLex contém muitos atributos associados com cada termo, e é difícil manter esta informação de forma estruturada e correta num sistema convecional enquanto RadLex cresce (...). Embora os documentos de texto e planilhas tenham sido em formato simples inicialmente para coletar os termos RadLex, eles não foram adequados para a distribuição de RadLex na Web ou para fazê-los acessíveis para as aplicações. Uma ferramenta adaptada às necessidades de desenvolvimento e

difusão da terminologia em radiologia é necessária (RUBIN, 2007⁵. *Tradução nossa*).

O projeto de atualização do Radlex teve início em 2005. De acordo com o site oficial da ferramenta, <http://bioportal.bioontology.org/ontologies>, o atual projeto é dirigido por um comitê de especialistas em informática, Radiologia e terminologias médicas, supervisionado pelo Comitê de Radiologia RS NA Informática. O site explica ainda que foram criados alguns comitês para a criação dos termos, na seguinte forma:

em 2005, seis comitês RadLex de sistema de órgãos foram formados em colaboração com mais de 30 organizações profissionais e de radiologia padrões, incluindo ACR, DICOM, e IHE. Em 2007, seis comissões adicionais foram recrutadas, cada uma focando uma modalidade de imagem específica (BIOPORTAL, 2011).

Esses comitês, cada um dentro de sua modalidade específica, definiram os termos para descrever os dispositivos, exames de imagem, e as etapas de procedimento realizado em radiologia, atividade que foi denominada de “Playbook RadLex” (BIOPORTAL, 2011), além de prestarem apoio ao comitê de direção, em contínuo desenvolvimento e cuidados com o RadLex.

Garantiu-se, com esses procedimentos, a cobertura abrangente de conteúdos, uma vez que a RSNA teve o apoio e a harmonização de outros vocabulários médicos: índice ACR, SNOMED-CT, CID9 (Classificação Internacional de Doenças, da Organização Mundial de Saúde), CPT (Terminologia Processual Atual), BrainInfo. Essa colaboração veio justamente da necessidade de integração de vocabulários existentes, com a finalidade de otimizar a pesquisa de imagens radiológicas.

Assim, o projeto Radlex, enquanto um vocabulário controlado da área da Radiologia, tem seus objetivos fixados na criação de um abrangente

⁵ Documento eletrônico

vocabulário de imagens médicas para capturar, indexar e recuperar uma variedade de informações sobre Radiologia. Para tal, foi desenvolvido adquirindo termos relevantes das subdisciplinas da radiologia: abdominal, cardiovascular, músculo-esquelético, pediátrico, torácico, neuroradiologia entre outras.

O objetivo final é distribuir o Radlex amplamente em formulários eletrônicos para que possam ser incorporadas aplicações, tais como documentos estruturados, criação de arquivos para ensino e dados de pesquisa para indexação (RUBIN, 2007).

5.5.2 A Plataforma Protégé no Radlex Ontology

A atual versão 3.5 do Radlex foi criada com padrões para distribuição na Web, utilizando a plataforma Protégé. Segundo Rubin (2007), essa ferramenta é bastante utilizada no meio científico para iniciar o desenvolvimento de uma Ontologia ou terminologia, por ser livre acesso (*open-source*). É de fácil manuseio, gratuito, suporta qualquer tecnologia utilizada para o desenvolvimento de terminologias e Ontologias, além de possuir uma gama extensa de aplicações de domínio.

O modelo Protégé para a representação da informação nas Ontologias e terminologias utiliza basicamente três *frames*: Classes, *Slots* e Facetas, além de quadros e axiomas (RUBIN, 2007). As classes são os termos que representam o domínio trabalhado. Os *Slots* são os espaços utilizados para descrever as propriedades ou atributos das classes. As Facetas descrevem as características dos *Slots*.

Além desses três frames básicos, o Protegé oferece outros que auxiliam na construção de Ontologias de maior complexidade, isto é, para a elaboração de vocabulários que tenham maior profundidade conceitual. No entanto, o Radlex utiliza apenas esses três *frames* básicos, porque suprem

satisfatoriamente as necessidades de Ontologias básicas, objetivas e claras, proporcionando fácil manuseio do vocabulário.

Outro fator que levou ao uso do Protégé foi o seu aspecto colaborativo, pois “suporta a edição colaborativa de Ontologias, permitindo aos usuários fazerem anotações nos componentes da Ontologia, bem como alterações na mesma” (RUBIN, 2007. *Tradução nossa*). A cada lançamento de uma nova versão, o vocabulário acrescenta novos termos e altera os antigos, com base nas sugestões dos usuários. Tais contribuições ao sistema são de natureza pragmática, uma vez que cada pesquisador sugere alterações de acordo com sua realidade contextual, ajudando a otimizar a semântica dos dados.

5.5.3 Estrutura do Radlex Ontology

O Radlex cobre o domínio do conhecimento da Radiologia, do campo da Medicina. Aborda os conteúdos da anatomia humana, questões patológicas, os dispositivos utilizados para capturar imagens, procedimentos e modalidades de imagens.

A forma de apresentação do Radlex é bastante clássica, partindo de uma lista de termos padrão e uma taxonomia que os organiza de forma sistemática, em categorias hierárquicas, onde os termos gerais são pais de termos mais específicos. Essa subordinação do termo específico ao termo pai é caracterizada na árvore hierárquica por relações formais denominadas de “*Is a*” e/ou “*subClassOf*”. Além disso, o vocabulário exibido na forma de taxonomia facilita a observação de todos os termos organizados hierarquicamente, o que facilita a detecção de inconsistências conceituais.

Esta captura de tela mostra a guia *Classes*, na qual o usuário cria, edita, e navega os termos RadLex. A taxonomia do Radlex é mostrada como uma árvore expansível do lado esquerdo. O painel direito exibe os detalhes de um termo RadLex

selecionado a partir da árvore à esquerda. Para cada classe selecionada, o Protégé exibe os atributos da classe (à direita), que os usuários podem editar. Por exemplo, o termo seio coronário foi selecionado (à esquerda) e é um ramo da área valvar supraaórtica. Outras guias no GUIA Protégé incluem: o *Slot* guia, para criação e edição de *slots*; na guia Formulários, para personalizar o *layout* de formulários de entrada de conhecimento (como a mostrada à direita); na guia instâncias para criar instâncias de classes e inserir valores *slot* particulares para aqueles casos, e qualquer outro para fins especiais "*Plug-in*" guia que o usuário pode querer usar. (RUBIN, 2007. *Tradução nossa*)

Figura 8: Interface Radlex Ontology

Fonte: Rubin, 2007

Conforme a figura 8, a taxonomia é mostrada à esquerda da tela de exibição do programa, tendo à direita a possibilidade de visualização da mesma em árvore hierárquica, mostrando todas as relações semânticas existentes no vocabulário, ou as informações de cada termo da taxonomia com o detalhamento devido. Qualquer termo selecionado na taxonomia aparecerá à direita do vídeo nas duas possibilidades: inserido na árvore ou os respectivos detalhamentos.

A árvore hierárquica é visualizada através de um plug-in chamado OntoViz, onde as classes são mostradas dentro de caixas. As relações entre elas são mostradas por meio de arcos que recebem um rótulo com o nome da relação, conforme figura 9. Esse *plug-in* é importante para a manutenção do vocabulário, visto que

O usuário pode selecionar as classes a serem exibidas no gráfico (painel inferior à esquerda), assim como os tipos de informações para mostrar, como subclasses, superclasses, e relações (painel superior esquerdo). Este paradigma de visualização é particularmente útil quando uma Ontologia contém mais de um tipo de relacionamento, já que a exibição em taxonomia (à esquerda), apenas mostra as classes da Ontologia de acordo com um relacionamento (*is-a*). (RUBIN, 2007. Tradução nossa)

Figura 9: Visualização do Radlex pela árvore hierárquica com o *plugin* OntoViz

Fonte: Rubin, 2007

As tipologias de classes e termos são constituídas por entidades específicas para a área biológica, dos tipos *continuada* e/ou *ocorrente*. Para Sowa (2000, p.67-77), continuantes, como o próprio nome sugere, são entidades que possuem existência contínua no tempo, preservando sua identidade através de mudanças ao longo de um intervalo de tempo prolongado. Ocorrentes são entidades que não têm identidade estável durante

qualquer intervalo de tempo ou mudança. Ou seja, possuem partes temporais, desdobram-se em fases sucessivas e existem somente em suas fases.

Figura 10: Continuanes e Ocorrentes

Fonte: Castilla, 2007

Castilla (2007, p.19) explica que as entidades continuantes se subdividem em duas: as continuantes independentes e as dependentes, onde as primeiras são relacionais com apenas um elemento de descrição (Castilla denomina de portador), incluindo “medida de temperatura, uma qualidade da forma ou uma doença”. Já as dependentes são relacionais com mais de um elemento de descrição, “como um ato de fotografar (ocorrentes) que, por exemplo, relaciona uma câmera (filme) a objetos fotografados, dois tipos de continuantes independentes”.

Considerando as reflexões de Sowa e Castilla, no caso da amostra deste estudo, as classes selecionadas das categorias **Radlex Entity** e **Radlex Descriptor**, a saber: “Anatomical Cluster” e “Object”, são basicamente entidades de natureza continuante dependentes, como mostram os detalhes da classe *Entity* abaixo.

Quadro 2: Detalhe da categoria “Radlex Entity”

Preferred Name (Preferred_name)	RadLex entity
ID	RID1
Full Id	http://purl.bioontology.org/ontology/RID/RID1
Comment	Currently all the Radlex entities are of the OBO Foundry ontological type "continuant".
Direct Superclasses	Radlex ontology entity
Direct Type	radlex metaclass
Name	RID1
Role	Concrete
Preferred Name	RadLex entity
Direct Subclasses	object procedure report component anatomical entity imaging observation Radlex non-anatomical set non-anatomical substance physiological condition assessment imaging modality property imaging procedure attribute procedure step report
Is A	Radlex ontology entity

Fonte: Bioportal, 2011

São entidades que expressam imagens derivadas de objetos concretos, os órgãos, necessitando para isso de vários elementos de descrição para a efetiva cobertura dos assuntos pertinentes às imagens.

Ainda em relação à tipologia dos termos, encontram-se dois elementos principais de representação no vocabulário: o termo preferencial, “aquele que os especialistas consideram ser a nomenclatura mais adequada da área”, e o termo sinônimo, que é a “forma alternativa de nomear os termos preferenciais” (BIOPORTAL, 2011).

Além dos termos básicos utilizados, citados acima, o vocabulário também inclui outros tipos de termos, os obsoletos. São aqueles que entraram no vocabulário, mas, por alguma razão, foram retirados definitivamente. Estes

formam uma lista à parte que pode ser consultada livremente pelos usuários.
Exemplo de termos obsoletos:

RID 12443

Comentário: foi usado como um nível de instância como exemplo de um procedimento; não necessário.

Nome preferencial para obsoleto: número de ecos de volta

Nome: RID 12443

Posição de termo: Retirado

Papel: concreto

RID 7265

Comentário: termo errôneo: o nervo radial superficial não tem um ramo lateral, mas um ramo dorsal que, por sua vez, tem ramos laterais e mediais.

Nome: RID 7265

Nome preferencial: ramo lateral de nervo radial superficial

Nome preferencial para obsoleto: ramo lateral de nervo radial superficial

Posição de termo: retirado

Papel: concreto

Os termos obsoletos ficam disponíveis por serem ainda utilizados nas práticas de pesquisas, sendo mantidos como entradas do vocabulário.

Além da tipologia de termos acima, o vocabulário também faz uso de atributos e de *slots*. Os atributos são conceitos que descrevem as informações auxiliares, tais como: definição do termo, mapeamento de termos relacionados em outras terminologias e informações de procedência do termo, ou seja, a fonte. Os *slots* são responsáveis pelas relações existentes entre os conceitos dos termos.

Como se pode observar, todos os elementos supracitados são utilizados para proporcionar uma interface amigável de manuseio, com capacidade para construir, preencher e ver as Ontologias de maneira simples.

A partir dessa explanação dos elementos de composição do Radlex Ontology, serão apresentadas as suas categorias e classes.

Métrica do Radlex Ontology

Número de classes: 34.895
 Número de indivíduos: 744
 Número de propriedades: 101
 Profundidade máxima: 23
 Número máximo de irmãos: 732
 Número médio de irmãos: 30
 Classes com uma subclasse única: 622
 Classes com mais de 25 subclasses: 59
 Classes sem definição: 34.879

Em 2008 possuía 1.326 termos
 Em 2011 possuía 7.413 termos
 Em 2012 possui aproximadamente 14.700 termos

Esses termos são classificados em duas categorias iniciais, a saber: **System Class** e **Radlex Ontology Entity**, que se desdobram da seguinte forma:

1: SYSTEM-CLASS

- :ANNOTATION
 - INSTANCE-ANNOTATION
- :CONSTRAINT
 - :PAL-CONSTRAINT
- :META-CLASS
 - :CLASS
 - :STANDARD-CLASS
 - radlex_metaclass
 - anatomy_metaclass
 - artery_metaclass
 - neuraxis_metaclass
 - nerve_metaclass
 - SetTermMetaclass
 - vein_metaclass
 - CoordTermMetaclass
 - ConjunctiveTermSetMetaclass
 - DisjunctiveTermSetMetaclass
 - imaging_sign_metaclass
 - obsolete_term_metaclass
 - pathophysiology_metaclass
 - :FACET
 - :STANDARD-FACET
 - :SLOT
 - :STANDARD-SLOT
 - radlex_attribute
- :RELATION
 - :DIRECTED-BINARY-RELATION

Nesta categoria estão organizadas todas as classes que dizem respeito ao funcionamento e organização do Radlex Ontology, ou seja, os termos que apresentam a forma e a terminologia das classes e subclasses utilizadas na construção da Ontologia.

2:RADLEX ONTOLOGY ENTITY

- CLASSE TERMOS OBSOLETOS RADLEX
- CLASSE NON-RADLEX TERM
- RADLEX DESCRIPTOR
- RADLEX ENTITY
- RADLEX TERM

Nessa outra categoria estão contidas as cinco classes que dizem respeito ao vocabulário dos assuntos pertinentes à área radiológica. O seu desdobramento resulta nos milhares de termos Radlex utilizados pelos pesquisadores da área. A seguir, as definições de cada classe da categoria:

Classe Non-Radlex term

Conforme definição do próprio vocabulário, esta classificação é usada para os nomes das entidades correspondentes em outras Ontologias, que consiste em mais do que uma sequência simples, como os termos Talairach. Esta classe não contém desdobramento de termos indexados (BIOPORTAL, 2011).

Classe Termos obsoletos radlex

RID's que existiam em uma versão anterior do Radlex mas não devem mais ser usados como parte do vocabulário. Esse recurso evita duplicações inadvertidas de termos, termos errôneos, etc. Se um termo é obsoleto, ou seja, não é mais usado na Medicina ou campos radiológicos, ele pode ser mantido na Radlex com uma nota no campo de comentário sobre seu *status* (BIOPORTAL, 2001).

Nesta classe estão ordenados 454 termos, quase todos identificados pelo seu código RID.

Classe descritor Radlex

Os termos desta classe são adjetivos, ao invés de sintagmas nominais como as entidades devem ser (BIOPORTAL, 2011). Possui 24 subclasses com 722 termos.

Classe Radlex Entity

Atualmente, todas as entidades Radlex são da OBO Fundação ontológica tipo "continuant", que é o caso dessa classe; é uma classe cujos termos correspondem a objetos físicos e não a processos. Essa é a classe mais extensa do Radlex, com 14 subclasses que se desdobram em aproximadamente 13.500 termos, a saber:

- 1- report component
- 2- non-anatomical substance
- 3- procedure step
- 4- object
- 5- anatomical entity
- 6- report
- 7- imaging procedure attribute
- 8- physiological condition
- 9- assessment
- 10- procedure
- 11- Radlex non-anatomical set
- 12- imaging observation
- 13- property
- 14- imaging modality

1 Report component

- [patient identifier](#)
- [report footer](#)
- [report header](#)
- [report section](#)
- [report subsection](#)

Esta classe contém os termos relacionados às únicas partes do relatório que podem conter textos ou termos que descrevem procedimentos e seus resultados. É uma classe pequena com cinco subclasses, contendo 18 termos.

2 [Non-anatomical substance](#)

- [chemical element](#)
- [pharmacological substance](#)
- [radioisotope](#)

Esta classe se desdobra, inicialmente, em três subclasses de assuntos, conforme exposto acima, e a partir das subclasses **[chemical element](#)** e **[pharmacological substance](#)**, relaciona 277 termos.

3 [Procedure step](#)

- [assess treatment response procedure step](#)
- [image interpretation procedure step](#)
- [imaging procedure step](#)
- [infusion](#)
- [interventional procedure step](#)
- [measurement procedure step](#)
- [medication administration procedure step](#)
- [therapy control procedure step](#)
- [treatment monitoring procedure step](#)
- [treatment planning procedure step](#)

Conforme linhas acima, a classe *procedure step* tem um desdobramento inicial considerável de 10 subclasses, dentre as quais seis se desdobram em aproximadamente 227 assuntos.

4 [Object](#)

- [foreign body](#)
- [medical object](#)

É uma classe que diz respeito aos assuntos não anatômicos, o que difere da maioria das classes do vocabulário (BIOPORTAL, 2011). Inicia com duas subclasses, no entanto, por meio delas são relacionados 371 termos radiológicos.

5 [Anatomical entity](#)

- [non-physical anatomical entity](#)
- [physical anatomical entity](#)
 - [anatomical set](#)
 - [immaterial anatomical entity](#)
 - [material anatomical entity](#)

Esta Classe diz respeito às entidades *continuants* (elementos atemporais) do organismo, delimitada pelo bom estado de um organismo ou é um atributo de sua organização estrutural. Exemplos: célula, coração, cabeça, cavidade peritoneal, ápice do pulmão, termo anatômico, plano sagital (BIOPORTAL, 2011). É a subclasse mais numerosa, composta de duas subclasses inicialmente, conforme exposto acima. No entanto, se desdobram, a partir da classe Physical anatomical entity, em mais três subclasses, através das quais são relacionados aproximadamente 9.390 termos, sendo:

aproximadamente 511 em **anatomical set** .
 aproximadamente 1.785 em **immaterial anatomical entity**
 aproximadamente 7.093 em **material anatomical entity**

6 [Report](#)

Essa classe diz respeito aos relatos descritos nas imagens radiológicas, ou seja, os laudos. Não contém desdobramento de termos.

7 [Imaging procedure attribute](#)

- [angiography attribute](#)
- [cross-sectional procedure attribute](#)
- [enhancement phase](#)
- [image processing attribute](#)
- [imaging acquisition](#)
- [imaging error](#)
- [interventional procedure attribute](#)
- [nuclear imaging attribute](#)
- [radiography procedure attribute](#)
- [ultrasound transducer element configuration](#)

Os termos desta classe são sintagmas nominais, relacionados a procedimentos de imagem. Nesta, inicialmente, são relacionados 10 assuntos. No entanto, cada assunto tem um desdobramento de dezenas de termos, totalizando aproximadamente 720 termos.

8 [Physiological condition](#)

- [non-pathological condition](#)
- [pathophysiologic entity](#)

Os assuntos relacionados aqui correspondem à condição fisiopatológica dos pacientes que deram origem às imagens radiológicas. É a segunda maior classe em número de termos. Como se pode perceber, a classe inicia com duas subclasses, mas tem um desdobramento de aproximadamente 1046 termos, todos organizados a partir da subclasse pathophysiologic entity.

9 [Assessment](#)

- [BI-RADS assessment](#)
- [intermediate concern](#)

É uma classe pequena que, por meio de duas subclasses iniciais, se desdobra em 13 termos, todos a partir da subclasse BI-RADS assessment.

10 [Procedure](#)

- [aspiration of tissue](#)
- [densitometry](#)
- [determine device status](#)
- [device manipulation](#)
- [device removal](#)
- [drainage](#)
- [embolization](#)
- [facet block](#)
- [follow-up procedure](#)
- [guidance](#)
- [imaging procedure](#)
- [limited imaging procedure](#)
- [localization](#)
- [marking for intervention](#)
- [nerve block](#)
- [nursing procedure](#)
- [pelvimetry](#)
- [placement](#)
- [quality control](#)
- [read outside image](#)
- [repair](#)
- [research](#)
- [screening](#)
- [shunt creation](#)
- [stereotaxis](#)
- [treatment](#)

Apesar de iniciar com 26 subclasses, não é uma classe grande em extensão de termos, possui aproximadamente 368 termos.

11 [Radlex non-anatomical set](#)

- [set of BI-RADS terms](#)
 - [set of BI-RADS mammo terms](#)
 - [set of BI-RADS MRI terms](#)
 - [set of BI-RADS ultrasound terms](#)

Um conjunto Radlex é uma classe que apresenta o número máximo de membros descontínuos da mesma classe (BIOPORTAL, 2011). É uma classe pequena que comporta apenas 4 termos que, ao nosso ver, garantem o agrupamento de objetos diversos do assunto BI-RADS.

12 [Imaging observation](#)

- [asymmetry](#)
- [destruction pattern](#)
- [echotexture](#)
- [enhancement](#)
- [enhancement pattern](#)
- [identifiable effect of mass on surrounding tissue](#)
- [imaging sign](#)
- [lesion boundary](#)
- [margin](#)
- [observed vascularity](#)
- [opacity](#)
- [symmetry](#)

Esta classe diz respeito aos assuntos relacionados às características observadas e relatadas pelos radiologistas durante a interpretação de uma imagem, que são os correspondentes visuais de doenças e processos patológicos dos pacientes que deram origem a estas imagens. No entanto, observações visuais não incluem as doenças reais ou processos que dão origem a elas (BIOPORTAL, 2011). Para tal, foram classificados inicialmente 12 assuntos, conforme exposto acima, através dos quais se desdobraram posteriormente em um total aproximado de 820 termos.

13 [Property](#)

- [clock face position](#)
- [count](#)
- [diagnosis](#)
- [diameter](#)
- [flow](#)
- [function](#)
- [image quality](#)
- [language of term](#)
- [measurement](#)
- [morphology](#)
- [patient maneuver](#)
- [patient position](#)
- [phase](#)
- [pressure](#)
- [slice thickness](#)
- [structure](#)
- [suction](#)
- [teaching descriptor](#)
- [temperature](#)
- [vascularity](#)
- [velocity](#)
- [viability](#)
- [volume](#)

Esta classe diz respeito às propriedades existentes nas imagens indexadas. Nela encontram-se, inicialmente, 23 classes que se desdobram em 175 termos.

14 [imaging modality](#)

- [dual energy xray absorptiometry](#)
- [fluoroscopy](#)
- [magnetic resonance imaging](#)
- [magnetic resonance spectroscopy](#)
- [nuclear medicine imaging](#)
- [panographic radiograph](#)
- [projection radiography](#)
- [spectroscopy](#)
- [tomography](#)
- [ultrasound](#)

Esta classe diz respeito à forma das imagens, isto é, ao meio utilizado para a imagem ser produzida, que resulta no seu formato. Aqui foram classificados inicialmente 10 classes, através dos quais foram classificados 52 termos.

5.5.4 Análise dos dados

Para efetuar a coleta de dados na Ontologia Radlex foi necessário realizar a análise prévia das classes, pois, conforme explicado no início esta seção, ela é uma Ontologia de grande porte que contém mais de 14 mil itens. Para os fins desta pesquisa, foi realizada a observação direta em cada categoria, visando à seleção de uma amostra que refletisse o conteúdo e a forma geral do vocabulário. Isso feito, chegou-se à seleção de quatro classes, sendo duas de cada uma das duas superclasses de maior representatividade do Radlex, formando assim uma amostra adequada para os objetivos da pesquisa. A sistematização dessa amostra ficou como segue:

Categoria Radlex Ontology Entity

➤ Class Radlex Descriptor

➤ Subclasses selecionadas do Radlex Descriptor:

- Image Observation Characteristic
- Pathophysiologic Process Modifier

➤ Class Radlex Entity

➤ Subclasses selecionadas do Radlex Entity:

- Anatomical Cluster
- Object

De posse das quatro classes selecionadas, procedeu-se à observação sistemática dos dados da coleta, conforme relato a seguir:

Todas as ocorrências continham os elementos básicos de representação, trazendo um nível de detalhamento com informações suficientes para a apresentação desses termos dentro de sua classe direta. Os detalhamentos básicos continham as seguintes informações:

- Identificação do termo com código na área radiológica – RID

- Localização do termo por URL
- Visualização da superclasse direta
- Identificação do tipo de termo – metaclass
- Função do conceito do termo – Role
- Relação Formal que liga o termo a outras classes
- A identificação de subdivisão se existisse – subclass

É importante salientar que todos os termos observados possuíam a função “concret”, conforme quadro 3:

Quadro 3: Detalhamento do termo “Anatomical plane”

Anatomical plane:

Preferred Name (Preferred_name) : anatomical plane

Synonyms (Synonym): fiat anatomical surface

ID : RID13250

Full Id : <http://purl.bioontology.org/ontology/RID/RID13250>

Direct Superclasses : immaterial anatomical entity

Direct Type : anatomy_metaclass

Name : RID13250

Role : Concrete

Definition : Anatomical surface which, as an imaginary plane, bisects an anatomical structure or an anatomical space. Examples: craniocervical plane, plane of pelvic inlet, temporal plane.

Fmaid :242982

Synonym : fiat anatomical surface

Direct Subclasses : anatomical transverse plane; sagittal plane; axial plane; imaging plane; multiple plane; anchored anatomical plane; oblique plane; coronal plane; longitudinal plane

Is A: anatomical surface

Fonte: Bioportal, 2011

Constatou-se que a natureza desses conceitos é da modalidade matéria/personalidade, sendo o todo e suas respectivas partes considerados objetos concretos. De acordo com Castillas (2007) e Sowa (2000), essa modalidade se enquadra em entidades Continuantes, conforme relatado no item sobre a estrutura do Radlex, onde se verificou que tal natureza do conceito advém do fato de serem imagens radiológicas, e não os elementos físicos a que essas imagens correspondem, ou seja, órgãos humanos e estruturas, processos biológicos e outros. Assim, se evidencia o porquê de não haver nenhum conceito derivado de processo ou atividade.

Subclasse: Image Observation Characteristic

Dos 221 termos existentes nesta classe apenas vinte e oito (28) possuíam definição e 18 continham comentários extras sobre o termo.

Em relação aos sinônimos, foram encontrados cento e setenta e seis (176) ocorrências nos termos analisados. Todos os sinônimos são termos equivalentes na língua alemã.

Além disso, foram identificados dois (02) pontos de acesso para outros vocabulários, com as respectivas ocorrências:

UMLs – 4 ocorrências

UMLs term – 4 ocorrências

O UMLS, ou Unified Medical Language System, “é um conjunto de arquivos e software que reúne muitos vocabulários padrão de saúde e biomedicina, que permite a interoperabilidade entre sistemas de computador” (UMLS Guia, 2012). Segundo o guia, esse vocabulário oferece uma listagem de termos que combina termos que vão desde informações médicas até nomes de remédios, ligando códigos entre médicos, farmácias, companhias de seguro e hospitais. No caso do Radlex, ele é considerado como mais uma fonte de busca para termos da área da Saúde.

No que diz respeito às relações formais, foi encontrada em dezenove (19) ocorrências a relação formal “member of” e, em todas as ocorrências, a relação formal “Is a”, ficando da seguinte forma a lista de relações formais encontradas nessa classe:

- Is a
- Member of

Ainda nesse item, além das relações formais, foi encontrada em duas (02) ocorrências a indicação “misspelling of term”, trazendo a maneira errada com que o termo foi escrito. A nosso ver, essa atividade reforça a eficiência

semântica do vocabulário, uma vez que, ao otimizar as questões de ordem sintática, clarifica-se o sentido.

Refletindo Sowa (2011⁶. *Tradução nossa*) nesta questão, observa-se que a expressão dos conceitos começa pelas palavras de uma linguagem, que “fornecem os devidos mecanismos gramaticais que combinam o seu estoque de palavras para expressar uma gama ilimitada de conceitos”; mas, partindo de palavras corretas e claras.

Subclasse: Pathophysiologic Process Modifier

Nesta subclasse foram identificados 55 termos, cuja análise mostrou que apenas uma (01) continha definição, o item 23, e nenhum deles comentários extras.

Em relação aos sinônimos, apenas cinco (05) termos não possuíam sinônimos, os itens 38, 39, 48, 55, 56 (ver anexo em cd-rom, classe pathophysiologic Process Modifier). O restante dos termos possuía expressões equivalentes em língua alemã.

No que diz respeito às relações formais, foi encontrada em todas as ocorrências uma única relação formal, a “/s a”, o que mostra que há apenas relações de hierarquia na classe.

Subclasse: Anatomical Cluster

Nesta subclasse foram identificados 184 termos, cuja observação encontrou na classe apenas cinquenta e cinco (55) continham definição, e nove (09) deles possuíam comentários extras.

Em relação aos sinônimos, noventa e oito (98) termos traziam equivalentes na língua alemã.

⁶ Documento eletrônico

Nesta classe foram encontrados dois (02) novos pontos de acesso em algumas ocorrências que remetem a vocabulários da área médica, são eles:

- FMAID – apareceu em 85 itens
- UMLs – apareceu em 22 itens

No que diz respeito às relações formais, foi encontrado um número considerável delas nessa classe, um total de doze (12), conforme lista abaixo:

- Is a
- Has part
- Part of
- Continuos with
- Has branch
- Member of
- Regional part of
- Has regional part
- Constitutional part of
- Has constitutional part
- Contained In
- Contained

Além das relações formais, foram encontrados também os seguintes pontos de acesso, que remetem para outros vocabulários da área da Radiologia e para outros termos:

- Linphatic Drainage
- Freesurfer
- Misspelling of term
- Talairach
- UMLs e UML terms

Subclasse: Object

Nos 155 termos observados foram encontrados os seguintes resultados:

- apenas dezenove (19) continham definição, e três (03) possuíam comentários extras.

- em relação aos sinônimos, cento e trinta e um (131) termos traziam sinônimos, todos em língua alemã.

- foram encontrados dois (02) novos pontos de acesso em alguns termos, os mesmo pontos de acesso a vocabulários da área Médica encontrados na classe Anatomical cluster, são eles:

- UMLs – apareceu em 12 itens
- UMLs term – apareceu em 12 itens

No que diz respeito às relações formais, apenas um (01) termo não possuía relação, no restante foi encontrado um número baixo nessa classe, um total de doze (02), conforme lista abaixo:

- Is a
- Member of

Resumindo, foram selecionados 615 termos na Ontologia Radlex, sendo identificadas 103 definições, 455 sinônimos e 13 relações formais.

5.5.5 Apresentação dos resultados da análise

Primeiramente foi realizada a análise de todas as relações existentes no vocabulário, procurando identificar as relações abordadas na literatura. Foram encontradas as seguintes relações:

Relações hierárquicas:

Na amostra, que compreendeu a análise de 615 termos, as relações com maior presença no vocabulário foram as **Relações Hierárquicas** do tipo genéricas, totalizando 99% das relações encontradas. Esse tipo de relação corresponde ao todo e suas partes, e no Radlex Ontology é utilizada com as nomenclaturas “SuperClassOf”, “SubClassOf” e “Is_a”, conforme exemplo abaixo.

Figura 11: Visualização do termo “Flow direction” na árvore hierárquica
 Fonte: dados da pesquisa em 2011

Quadro 4: Detalhamento do termo “Anatomical compartment”

2 Preferred Name (Preferred_name)	anatomical compartment
ID	RID29528
Full Id	http://purl.bioontology.org/ontology/RID/RID29528
Direct Superclasses	anatomical cluster
Direct Type	anatomy metaclass
Name	RID29528
Role	Concrete
Definition	Anatomical cluster which has as its parts one or more anatomical structures surrounding an anatomical compartment space and the contents of that space. Examples: mediastinal part of chest, anterior compartment of forearm, palm of hand(compartment), intercostal space, vertebral canal
Fmaid	9647
Preferred Name	anatomical compartment

Direct Subclasses	compartment of trunk
Is A	anatomical cluster

Fonte: Bioportal, 2011

Apenas um termo não possuía a relação *Is_a*, a ocorrência **26**, apenas as relações SubClass e SuperClass. Todas as demais possuíam esse tipo de relação como base das relações formais. Em alguns casos observou-se a apresentação de outras formas de relação, mas a base foi a hierarquia.

Essa constatação vem confirmar o que a literatura revela sobre as relações de maior ocorrência nas Ontologias. Conforme Smith, essas relações se configuram como fundamentais nas Ontologias, ou seja,

estabeleceram-se como fundamental para Ontologias atuais. Eles têm um papel central em quase todas as Ontologias de domínio, incluindo o modelo fundacional de Anatomia (FMA), GO e outras Ontologias em OBO, bem como em influentes Ontologias de nível superior como Dolce e em digitalizados recursos lexicais tais como WordNet (SMITH, 2005, p.5. *Tradução nossa*).

Guarino (1998) discute em seu artigo justamente essa sobrecarga da relação *Is_a* nos vocabulários das Ontologias, afirmando que esse tipo de relação cobre quase 100% das relações existentes em uma Ontologia, mas que, infelizmente, não contribuem para dar consistência aos vocabulários. Segundo o autor, as inconsistências se devem à falta de coerência entre a relação *is_a* e o tipo que ela representa, gerando, muitas vezes, “confusão de sentidos, redução de sentidos, super-generalização, links de tipos de função duvidosos e confusão de funções taxonômicas” (GUARINO, 1998 apud SALES, 2006, p.65).

O uso predominante de relações hierárquicas nas Ontologias, gera preocupação na área da recuperação da informação. Os instrumentos de tratamento, recuperação e descoberta de informações serão tanto mais eficazes quanto mais variadas forem as relações entre os termos. Pode-se afirmar que o predomínio das relações hierárquicas compromete as funções propostas para as ontologias. De fato, a ausência de relações semânticas

(sinonímia) ontológicas (relações temporais e espaciais), reduzem o poder de operação das ontologias por ficar comprometida de maneira significativa a produção de sentido dos conceitos, impossibilitando, desse modo, a efetiva recuperação de conteúdos por parte dos usuários.

Considerando que a principal função das ferramentas ontológicas é a identificação do contexto nos conteúdos, e que esse contexto se dá de maneira geral pelas relações existentes nas Ontologias, observa-se de início uma certa fragilidade do Radlex. As questões de contexto para a efetivação da significação são fundamentais nos sistemas de informação, principalmente em sistemas mediados por tecnologias de informação, onde as perguntas e respostas que orientam o tratamento da informação com vistas à inferência lógica, precisam ser estabelecidas de forma a evitar ambiguidades de sentidos.

- Relações de equivalência:

Essas relações são encontradas no Radlex com a nomenclatura “**synonyms**”, onde se verifica um nome alternativo para o termo preferencial. Segundo o Bioportal, esses sinônimos, apesar de não terem sido escolhidos como nomes preferenciais, são utilizados pelos especialistas da área de radiologia, por isso, colocados como entradas de acesso dos termos.

Quadro 5: Detalhamento do termo “Blinded interpretation”

Preferred Name (Preferred_name) : blinded interpretation
Synonyms (Synonym) : geblindete Interpretation
 ID : RID13058
 Full Id : <http://purl.bioontology.org/ontology/RID/RID13058>
 Direct Superclasses : [image interpretation mode](#)
 Direct Type : [radlex metaclass](#)
 Name : RID13058
 Role : Concrete
 Is A : [image interpretation mode](#)

Fonte: Bioportal, 2011.

Foi constatado que 80% dos 615 termos analisados possuem sinônimos, se constituindo num total de 455 termos com nomes alternativos. É importante ressaltar que 100% desses sinônimos se encontram na língua estrangeira, mais especificamente no alemão, e apenas 01 ocorrência saiu desse padrão, possuindo, além do equivalente em alemão, outro em italiano.

Grafico 1: Porcentagem de relações formais, definições e sinônimos

Fonte: dados da pesquisa em Nov. de 2011

Acredita-se que esse padrão da língua alemã na área radiológica se deve ao fato de a radiologia ter seu berço na Alemanha, onde os estudos continuam avançando e, conseqüentemente, procurados pelos especialistas da área, como afirma a Sociedade Paulista de Radiologia e Diagnóstico por Imagem.

A história da Radiologia começou em 1895 com a descoberta experimental dos raios-X pelo físico alemão Wilhelm Conrad Roentgen. À época as aplicações médicas desta descoberta revolucionaram a medicina, pois havia se tornado possível a visão do interior dos pacientes. Com o passar dos anos, este método evoluiu e assumiu uma abrangência universal na pesquisa diagnóstica do ser humano. (Sociedade Paulista de Radiologia e Diagnóstico por Imagem, 2012)

Os sinônimos, no caso do Radlex, são considerados de grande utilidade para viabilizar a comunicação do vocabulário com os usuários, uma vez que são empregados de acordo com a necessidade dos pesquisadores. A inclusão de sinônimos contribui para a expressividade semântica da ferramenta. De fato, os sinônimos e remissivas estabelecem ligação entre o conteúdo tratado e os termos empregados nas buscas.

- Relações Formais

Nesse item foram analisadas as relações formais existentes no vocabulário, observando como as mesmas são utilizadas para representar o contexto que liga os conceitos nas relações categoriais. Essa atividade é fundamental para o tratamento da informação, uma vez que ela é uma das peças fundamentais para o desenho da expressão conceitual de uma Ontologia.

Assim, entre os 615 termos constitutivos do *corpus*, foram encontrados 12 tipos diferentes de relações formais, a saber:

Is_a
 Part_Of
 Has_Part
 Continuos_With
 Has_Branch
 Member_Of
 Regional_Part of
 Has_Regional Part
 Constitutional_Part Of
 Has_Constituional Part
 Contained_In
 Contained

Algumas dessas relações já mencionadas anteriormente, como a *Is_a*, são encontradas na literatura, como é o caso das *Is_a*, *Part_Has_part* e *Contained_in*. As outras relações foram criadas especificamente para suprir as necessidades conceituais do Radlex Ontology.

No entanto, podem-se constatar os seguintes aspectos das relações encontradas:

- a existência das relações fundamentais na construção de uma Ontologia: *is_a* e *part_of*, as quais sugerem no vocabulário a hierarquização dos assuntos em parte e todo. Aliás, uma atividade essencial em qualquer vocabulário, para cumprir os objetivos terminológicos, de informar literalmente a relação expressa entre as categorias de assuntos organizados.
- a existência das relações inversas, tão sugerida na literatura de organização da informação:

Part_of* sua inversão *Has_part

Regional_Part of* sua inversão *Has_Regional Part

Constitutional_Part Of* sua inversão *Has_Constitutuional Part

Foram encontrados *no corpus* 12 tipos diferentes de relações formais, a saber:

Is_a

Part_Of

Has_Part

Continuos_With

Has_Branch

Member_Of

Regional_Part of

Has_Regional Part

Constitutional_Part Of

Has_Constitutuional Part

Contained_In

Contained

Algumas dessas relações, já mencionadas anteriormente, são encontradas na literatura, como é o caso das *Is_a*, *Part_Of*, *Has_part* e *Contained_in*, *Has_branch*, *Constitutional_Part Of*, *Continuos_With*. As outras são relações que foram devidamente criadas para suprir as necessidades conceituais do Radlex Ontology.

Constatou-se, ainda, que se a totalidade de relações formais foi considerada pequena, as relações semânticas podem ser consideradas quase inexistentes, pois apresentaram apenas 2 relações, as hierárquicas e as associativas, sendo explicitadas da seguinte forma no Radlex:

- Relações formais que descrevem as relações hierárquicas:

- Is_a*
- Part_Of*
- Has_Part*
- Has_Branch*
- Member_Of*
- Regional_Part of*
- Has_Regional Part*
- Constitutional_Part Of*
- Has_Constitutuional Part*

- Relações formais que descrevem as relações associativas:

- Continuos_With*
- Contained_In*

6 Considerações finais

Os estudos sobre Ontologias promovem a aproximação e a adaptação de normas de representação e tratamento da informação aos novos canais baseados em tecnologias da informação. A presente pesquisa procurou refletir teoricamente sobre os pressupostos que orientam os estudos de semântica para a recuperação da informação no ambiente da Web.

Primeiramente, o estudo mostrou a importância das ontologias para o tratamento da informação no ambiente virtual, haja vista que traz, em sua estrutura, aplicativos automatizados relevantes para a construção de linguagens documentárias pertinentes às necessidades do cenário atual da informação, fato que não é encontrado em instrumentos convencionais, como os sistemas de classificação e os tesauros. Entre os novos tipos de aplicativos podem-se destacar mecanismos que propiciam: o reuso de terminologias em diferentes plataformas, a interoperabilidade entre diferentes sistemas, a visualização de todas as relações semânticas do vocabulário em uma árvore hierárquica num simples toque, que facilita a identificação de inconsistências semânticas dos termos, uma vez que se consegue observar as relações de maneira generalizada.

Embora as questões computacionais sejam relevantes, as terminologias são essenciais para os fins de recuperação da informação. Essa afirmativa ficou ainda mais evidente quando foram processados os dados desta pesquisa, cujos resultados deixam claro que não basta a existência de aplicativos automatizados potentes para o bom resultado desses instrumentos, sobretudo porque os mesmos se constituem de linguagens que controlam não apenas a sintaxe (aspecto linguístico bastante trabalhado nos canais computacionais),

mas também os conteúdos semânticos de domínios do conhecimento, elemento relacionado com o feedback oriundo da relação entre os usuários e o sistema de conceitos desses instrumentos. Por esse motivo, considera-se que esta pesquisa sobre as Ontologias favorece as atividades da área de Organização da Informação, considerando-se as questões relativas à produção de sentido que foram aqui exploradas.

No caso do Radlex Ontology, alguns aspectos observados foram bastante marcantes para compreender a representação de conteúdos, alguns deles positivos e outros negativos.

Logo de início, se observou que, para o processo de representação, que consiste na construção de significados, o vocabulário Radlex traz todos os elementos básicos pertinentes ao ato de representar, elementos esses previstos pelas diretrizes da norma internacional ANSI/2005, incluindo as definições, encontradas nos termos que necessitavam de uma notificação específica para sanar sentidos dúbios de conceitos; as notas de escopo, denominadas no Radlex de “comentários”, onde também se esclarecia alguma especificação informacional do termo; as relações lógico-semânticas, incluindo as hierárquicas, as equivalentes e as associativas; e a formatação dos termos, incluindo os preferenciais e os não-preferenciais.

No que se refere às questões terminológicas, o Radlex possui alguns elementos que fortalecem a produção de sentido num processo genuinamente pragmático, no sentido de estabelecer uma relação forte entre o sistema de conceito do vocabulário e as necessidades dos pesquisadores e técnicos da área da radiologia. Um desses elementos que, a nosso ver, é de extrema importância para esse processo é o canal de comunicação que o Radlex tem com os pesquisadores, por meio do qual tem-se a oportunidade de dar valiosas contribuições terminológicas ao vocabulário, sugerindo novas nomenclaturas para os termos existentes no vocabulário, novos termos, novos assuntos, novas composições de conceitos, entre outras ações que ajudam a fortalecer o sistema. De acordo com site do programa, essas contribuições têm influência direta na reformulação dos conceitos e dos termos, a cada nova edição

lançada, sugerindo uma efetiva atividade pragmática dinâmica em todo o processo de construção de sentidos.

Peirce (1977) afirma que o processo de representação é uma relação “genuinamente triádica”, onde o signo e o objeto não atuam sozinhos, tendo automaticamente a presença do elemento mediador, o interpretante, colocado como um signo externo ao processo que, com natureza de lei, evoca definitivamente os significados necessários para a construção de sentidos.

Nesse caso, pode-se dizer que os termos e os conceitos sugeridos pelos pesquisadores são os elementos mediadores do processo de representação, tal como proposto por Peirce. São conceitos externos às atividades de representação, oriundos da vivência e da bagagem cultural e intelectual dos usuários do sistema, que funcionam como o terceiro elemento do processo de representação, os criadores dos interpretantes. Ora, o próprio Peirce coloca, de maneira bastante didática, que é possível esclarecer o sentido real de um signo apontando um objeto, fazendo uma mímica ou até mesmo lendo um discurso completo; são atitudes que auxiliam a entender o signo trabalhado.

Assim, quando observada a formalização das relações e as próprias definições dos termos, é possível visualizar, de forma concreta, a presença da teoria do interpretante de Peirce (1977). Com efeito, as relações e as próprias definições dos termos decorrem de atividades pragmáticas que compõem o terceiro elemento, aquele que dá significado aos conceitos. Nesse sentido, refletindo a respeito do processo de significação peirceano, Eco (2007) afirma que não basta substituir o objeto por um signo, ao contrário, são necessárias informações do tipo “que me fazem saber algo mais”. Assim, na formulação de alguns termos do vocabulário, tais como o termo “*anatomical cluster*”, foi necessário algo mais para obter uma especificação conceitual mais precisa, para tirar de foco toda possibilidade de duplicação de sentido. Nesse termo pode-se visualizar uma definição bastante detalhada para compor o conceito, conforme quadro abaixo, além de outras informações, como o sinônimo e a remissiva para outro vocabulário.

Quadro 6: Detalhes do termo “Anatomical Cluster”

1 Preferred Name (Preferred_name)	anatomical cluster
Synonyms (Synonym)	anatomischer Block
ID	RID13290
Full Id	http://purl.bioontology.org/ontology/RID/RID13290
Direct Superclasses	anatomical structure
Direct Type	anatomy metaclass
Name	RID13290
Role	Concrete
Definition	Anatomical structure, which has as its parts a heterogeneous collection of organs, organ parts, cells, cell parts or body part subdivisions that are adjacent to, or continuous with one another; does not constitute a cell part, cell, tissue, organ, organ system or organ system subdivision, cardinal body part, body part subdivision or anatomical junction. Examples: joint, adnexa of uterus, root of lung, renal pedicle, back.
Fmaid	49443
Preferred Name	anatomical cluster
Synonym	anatomischer Block
Direct Subclasses	meningeal cluster heterogeneous cluster organ cluster organ part cluster anatomical conduit scalp carpal bridge region soft palate cell part cluster body part subdivision cluster bony palate vestigial embryonic structure anatomical junction pelvic floor inner ear upper esophageal sphincter renal collecting system compartment subdivision palate

	anatomical compartment middle ear tongue gestational structure
Is A	anatomical structure

Fontes: Bioportal, 2011

O que se pode perceber é que essas definições são formuladas dentro de um processo pragmático, de acordo com contextos que refletem o dia a dia dos pesquisadores da área de radiologia, em consonância com a realidade de quem recebe o signo.

No que diz respeito às relações semânticas do vocabulário, a primeira questão observada foi a existência da relação de equivalência, que consiste na ligação entre os termos não preferidos e os termos preferidos do sistema de conceitos. Essa relação se configura no vocabulário por meio de sinônimos que remetem aos termos preferidos. Esses sinônimos (termos não preferidos) no Radlex, são na sua totalidade termos equivalentes na língua alemã (tendo 2 na língua Italiana), se constituindo em novas entradas para os termos Radlex.

Considerando que 80% das ocorrências trazem sinônimos como termos alternativos para a efetiva utilização do Radlex, pode-se dizer que eles cobrem uma parte significativa dessa ontologia, o que beneficia sobremaneira o processo semiótico, haja vista, que a representação em vários contextos, segundo Kobashi (1996), caracteriza uma função sígnica. Pode-se afirmar, a partir disso, que a caracterização contextual dos signos é um processo de eficácia comprovada, uma vez que apresenta algo (o conteúdo) por meio de diferentes signos ou símbolos que abrangem características sintáticas e semânticas imersas em elementos culturais como a própria língua. Esse tipo de abordagem representacional garante a interpretação correta (desejada e esperada), no caso das Ontologias de domínios específicos.

Por outro lado, embora existam muitas definições e sinônimos que remetem aos termos equivalentes de outro idioma, não se encontrou nenhuma

ocorrência que apresentasse sinônimos com variantes lexicais no próprio idioma, uma operação prevista nas relações de equivalência das normas ANSI/2005, com a finalidade de promover uma significação mais precisa dos termos de um vocabulário.

Um exemplo de como isso poderia ocorrer no Radlex é o termo “*lungs*” (Pulmão), encontrado na subclasse “*Subdivision of organ system*” (conforme quadro 7).

Quadro 7: Detalhes do termo “Lungs”

153 Preferred Name	Lungs
ID	RID13437
Full Id	http://purl.bioontology.org/ontology/RID/RID13437
Direct Superclasses	subdivision of organ system
Direct Type	anatomy metaclass
Name	RID13437
Role	Concrete
Definition	union of left lung and right lung
Part Of	thorax
Preferred Name	Lungs
Is A	subdivision of organ system

Fonte: dados coletados na pesquisa em 2011

A nosso ver, o termo poderia receber outros termos sinônimos para melhor eficácia da ontologia, tais como: *bellows* e *windbag*, palavras em inglês cujos significados são sinônimos ou quase-sinônimos de *Lungs*.

Nesse sentido, relembramos Armengaud (2006, p.79), sobre a abordagem contextual: “os termos certos para cada lugar e identidade, (...) ajudaria na interação entre o usuário e a informação tratada (ou signos carregados de interpretações), respeitando as pressuposições que se tornaram comuns aos mesmos através de suas realidades culturais”.

Seguindo na análise das relações, constatou-se também a existência das Relações Hierárquicas, do tipo genérica, que totalizou 100% das ocorrências. Esse resultado levanta preocupações em relação à semântica do Radlex.

Primeiro, verificou-se que os termos desta ontologia estão relacionados majoritariamente de forma hierárquica - relações *todo e parte/gênero e espécie*-, deixando de lado, ou ocultando, outros tipos de conceitos e relações que contribuiriam para uma melhor constituição conceitual do Radlex. Nesse sentido, pode-se dizer que as questões semânticas estão prejudicadas, o que pode comprometer a precisão dos sentidos dos constituintes da ontologia.

Essa constatação vem confirmar o que a literatura revela sobre as relações de maior ocorrência nas Ontologias. Conforme Smith, as relações genéricas são dominantes nas Ontologias:

“estabeleceram-se como fundamental para as Ontologias atuais. Elas têm um papel central em quase todas as Ontologias de domínio, incluindo o modelo fundacional de Anatomia (FMA), GO e outras Ontologias em OBO, bem como em influentes Ontologias de nível superior como Dolce e em digitalizados recursos lexicais tais como WordNet” (SMITH, 2005, p.5. *Tradução nossa*).

Retomamos Guarino (1998) que discute em seu artigo justamente a sobrecarga da relação *Is_a* nos vocabulários das Ontologias, afirmando que esse tipo de relação cobre quase 100% das relações existentes em uma Ontologia, mas que, infelizmente, não contribuem de maneira consistente para promover a qualidade desses instrumentos. Segundo o autor, essa inconsistência se deve à falta de coerência entre a relação *Is_a* e o tipo a que ela representa muitas vezes, gerando “confusão de sentidos, redução de sentidos, super-generalização, links de tipos de função duvidosos e confusão de funções taxonômicas” (GUARINO, 1998 apud SALES, 2006, p.65).

Considerando que uma das principais características das ferramentas ontológicas é identificar do contexto nos conteúdos, e que esse contexto se dá de maneira geral pelas relações existentes nas Ontologias, a pouca variedade de relações semânticas, já que nelas predominam as relações lógicas (genéricas), pode-se afirmar que o Radlex fica comprometido quanto à sua

capacidade de significar para representar informação ou conteúdos. As questões de contexto para a efetivação da significação são de extrema importância nos sistemas de informação para que não haja duplicidade e ambiguidade na formação de sentidos. Em relação a isso, Gonzáles de Gómez (2007, p.4) explica que é preciso nessas atividades esclarecer “qual é o caso em que a informação é o caso”, para que haja a “ação Informativa” tão necessária nos processos de significação das mensagens recebidas.

Reportando essa situação ao processo de semiose peirceana, pode-se dizer que o uso de relações que apresentem contextos dentro de limites adequados, aproximados, serve de base para que o interpretante, enquanto novo signo do processo de formação de sentido, se ajuste e desenvolva efetivamente uma significância real existente entre o signo de representação e o objeto representado.

No tocante às relações formais, com base nos dados analisados, constatou-se que apesar de ser um vocabulário de grande porte, contendo mais de 14.000 termos, as relações formais refletem as relações semânticas, ou seja, ambas são limitadas.

No que diz respeito às formas, constatou-se que das 12 encontradas na Ontologia, 5 são relações que estão estabelecidas na literatura; as outras são relações específicas nascidas das necessidades conceituais do vocabulário. Isso comprova o que a literatura afirma sobre a falta de diretrizes para a criação e utilização das relações formais nas ontologias. Em estudos nessa área (SALES, 2006), constata-se que cada ontologia cria as relações formais de acordo com as suas necessidades, estabelecendo conexões conceituais dentro de diretrizes que ela mesma formula, comprometendo, de certa forma, as condições de interoperabilidade.

Além disso, verificou-se também que, além de poucas relações formais, existe a supremacia das relações do tipo fundamental (relembrando: *Is_a*; *Part_of*; *Has_Part*; *Has_branch*; *Member_of*; *Regional_Part of*; *Has_Regional_Part*; *Constitutional_Part of*; *Has_Constitutional_Part*), as quais sinalizam no vocabulário a hierarquização dos assuntos em *gênero/espécie* e

parte/todo. Por outro lado, essa constatação mais uma vez comprova a fragilidade da estrutura conceitual da ontologia Radlex, sobretudo quando a literatura e as normas prevêm a necessidade de maior variedade de relações formais e semânticas para a construção de significados de vocabulários controlados. Ora, se o signo não substitui o objeto, muito pelo contrário, é determinado por ele, indicando a idéia que ele mesmo produz, então, um vocabulário com a riqueza de objetos concretos constituídos por imagens (elementos por si só soberanos em significações), como é o caso do Radlex, não poderia reduzir e simplificar as relações que a constituem. A supremacia da significação do signo sobre o objeto não deixa dúvidas no processo semiótico: sem a significação devida, o objeto (conteúdos) dificilmente será recuperado de forma eficaz, sobretudo quando esse objeto está armazenado em meio a uma infinidade de terabytes de informação.

Nesse sentido, a teoria peirceana, que se preocupa diretamente com as questões terminológicas, ressalta que o signo interpretante constitui um significado primeiro e externo ao signo da representação, o que leva à conclusão de que não existe nenhum modo de se entender um signo a não ser pelo seu interpretante. A construção de sentidos, fiéis aos signos e conteúdos representados, no âmbito das ontologias, será efetivamente contemplada quando a sua estrutura conceitual refletir os preceitos da semiótica.

Portanto, considerando a amostragem e a análise dos dados, evidenciou-se que apesar da grande divulgação e utilização desse instrumento semântico, na atualidade, surgem, por vezes, pseudo ontologias. Os dispositivos, embora se auto-definam como instrumentos ontológicos, apresentam a estrutura física e semântica de taxonomias. Em outras palavras, o Radlex Ontology, apesar do título, configura-se, de fato, como uma taxonomia, que possui uma lista de termos organizados hierarquicamente, com relações semânticas simples do tipo todo-parte, apresentando em sua estrutura as mesmas limitações semânticas de ferramentas tradicionais de tratamento e recuperação de informação: taxonomias e tesauros. Observou-se, além disso, que existem limites nas linguagens lógicas para expressar significados, o que

leva a crer que os aspectos semânticos investidos na Ontologia analisada, são tão limitados quanto os dos instrumentos tradicionais.

Assim, a teoria peirceana, incluindo sua trilogia de signos, as relações de significação, a semiose, suas reflexões sobre a representação e, sobretudo, a teoria do interpretante, mostrou ser um aporte teórico relevante para observar as relações presentes nas Ontologias.

O estudo das Ontologias é complexo, requer diferentes aportes teóricos e procedimentos. Nesta pesquisa, os resultados referem-se à análise de uma Ontologia por meio de categorias e conceitos advindos da teoria peirceana, e da lógica. Novos estudos, sob diferentes perspectivas e métodos são necessários para consolidar conhecimentos sobre a construção e avaliação de ontologias para sistemas de grande porte. É o que pretendemos desenvolver em estudos futuros.

REFERÊNCIAS

ALLWOOD, Jeans; ANDERSSON, Lars-Gunnar; DAHL, Östen. **Lógica para Linguístas**. Madri: Paraninfo, 1981.

ALMEIDA, M. B.; MOURA, M.A; CARDOSO, A. M. P.; CENDON, B. V. Uma Iniciativa Interinstitucional para Construção de Ontologia sobre Ciência da Informação: Visão Geral do Projeto P.O.I.S. **Rev. Eletr. Bibliotecon. Cin. Inf.**, Florianópolis, n.19, 1º sem., 2005.

ALMEIDA, Mauricio B.; BAX, Marcello P. Uma visão geral sobre Ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Rev. Cin. Inf.**, Brasília, v.32, n.03, p. 7-20, set/dez. 2003.

ANSI.ISO Z39.19-2005: Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Maryland, U.S.A.: National Information Standards Organization, 2005.

ARAÚJO, C. A. A. **A Ciência da informação como ciência social**. *Ciência da informação*. v.32, n.3, Brasília set./dez. 2003.

ARMENGAUD, F. **A pragmática**. São Paulo: Parábola, 2006.

AZEVEDO NETTO, Carlos Xavier de. Signo, sinal, informação: as relações de construção e transferência de significados. **Informação & Sociedade: estudos**, v.12, n.2, 2002.

BENTES PINTO, V. **La representation des connaissances dans le contexte de la documentation technique**: proposition d'un modèle d'indexation. Grenoble, 1999. (Thèse de doctorat, Université Stendhal Grenoble-3).

Bioportal. Disponível em <http://www.bioportal.bioontology.org/>. Acesso em 2011/2012.

BOUGNOUX, Daniel (org.). **Science de l'information et de la communication**. Paris: Larousse, 1993.

BREITMAN, Karin. **Web semantica**: a internet do futuro. Rio de Janeiro: LCT, 2005.

CABRERA, Julio; SILVA FILHO, Otavio Leopoldino da. **Inferências Lexicais e Interpretação de Redes de predicados**. Brasília: Ed. UNB: Finatec, 2007.

CAPURRO, Rafael. Epistemologia e Ciência da Informação. **V Encontro Nacional de Pesquisa em Ciência da Informação**. Belo Horizonte (Brasil), 10 de Novembro de 2003. Disponível em <http://www.capurro.de/enancib_p.htm> Acesso: 06/09/2007

CAMPOS, J. A. G. **Análise conceitual sobre as relações semânticas em Ciência da Informação**: contribuições para o desenvolvimento de Ontologias. 146 f. Dissertação (Mestrado em Ciência da Informação) - Escola da Ciência da Informação , Universidade Federal de Minas Gerais , Belo Horizonte, 2009.

CAMPOS, Maria Luiza de Almeida. **A organização de unidades de conhecimento: o modelo conceitual como espaço comunicacional para a realização da autoria**. 2001. Tese (Doutorado em Ciência da Informação) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

CARDOSO, J. B. F. **A Semiótica do Cenário Televisivo**. São Paulo: Annablume, 2008.

CASTILLA, André Coutinho. **Instrumento de investigação clínico-epidemiológica em cardiologia fundamentado no processamento natural de linguagem natural**. 2007. Tese (Doutorado em Medicina) – Universidade de São Paulo, São Paulo, 2007.

CHOMSKY, Noam. **Reflexões sobre a Linguagem**. São Paulo: Cutrix, 1980.

DAVIES, J.; Fensel, D.; Van Harmelen, F. **Towards The Semantic Web: Ontology-Driven Knowledge Management**. John Wiley & Sons Ltd, 2003. (e-book)

ECO, U. **Quase a Mesma Coisa**. Rio de Janeiro: Record, 2007.

FALBO, Ricardo A., et al. Ontologias e Ambientes de Desenvolvimento de software Semânticos. Disponível em: <<http://www.inf.ufes.br/~falbo/download/pub/2004-JIISIC-1.pdf>> Acesso em: 08/09/2007

FERNANDES, Joliza C.; BENTES PINTO, V ; AZEVEDO NETTO, C. X. . Representação Indexal na WEB: Estudos do Sintagma. **Rev. Informação & Sociedade**. Estudos, João Pessoa - Pb, v. V.15, n. 2, p. 1-20, 2006.

FERREIRA, F. J. I. No paraíso, sem convicção... (uma explicação do programa de Hilbert)." In: **Matemática e Cultura II** (Furtado Coelho, ed.), Centro Nacional de Cultura and SPB Editores, Lisbon, 1995. p. 87-121.

FOUCAULT, Paul-Michel. **A ordem do discurso**. São Paulo: Loyola, 1996.

FROHMANN, Bernd. **Taking information policy beyond information science: applying actor network theory.** Paper apresentado no 23rd Annual Conference of Canadian Association for Information Science. Edmonton, Alberta, jun. 1995. Disponível em <http://www.fims.uwo.ca/people/faculty/frohm/actor.htm> Acesso em 08.junho.2012.

GONZÁLEZ de GOMÉZ, M^a Nélida. As relações entre ciência, Estado e sociedade: um domínio de visibilidade para as questões de informação. **Ci. Inf.**, v.32, n.1, p.60-76, Brasília, Jan/Abr, 2003.

GONZALEZ de GOMEZ, M^a. Nélida. Para una reflexión epistemológica sobre la ciencia de la información. **Signo y Pensamiento**, v. 50, p. 10-22, 2007.

GRUBER, Thomas R. A translation approach to portable ontologies. **Knowledge Acquisition**, v. 5, n. 2, 1993, p. 199-220

GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing. In: **International Journal Human-Computer Studies (IJHCS)**, 43, 1993. Proceedings... Padova: Ladsed, p.907-928, 1993.

GRUBER, Thomas R. **What is an ontology?** Disponível em <http://www.ksl.stanford.edu/ksl/what-is-an-ontology.html> Acesso em: 14/09/2007.

GRUBER, Thomas R., OLSEN, Greg R., RUNKEL, J. **The configuration design ontologies and the VT elevator domain theory.** *International Journal of Human-Computer Studies*. Volume 44 , Issue 3-4 March/April 1996.

GUARINO, N. Formal Ontology and Information Systems. in: N. Guarino, (Ed.) **Formal Ontology in Information Systems**. Amsterdam, Netherlands: IOS Press, 1998. p. 3-15.

GUARINO, N. Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration. in: M. Pazienza, (Ed.) **Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology, International Summer School, SCIE-97**, Frascati, Italy, p. 139-170, 1997b.

GUARINO, N. Understanding, building and using ontologies. **International Journal of Human-Computer Studies**, v. 46, n. 2-3, p. 293-310, Feb, 1997a.

IBAÑOS, A. M. T.; SILVEIRA, J.R.C. **Na Interface Semântica/Pragmática: programa de pesquisa em lógica e linguagem natural.** Porto Alegre: EDIPUCRS, 2002.

JOÃO BRANQUINHO, J.; MURCHO, D.; GOMES, N.G. **Enciclopédia de Termos Lógico-filosóficos.** São Paulo: Martins Fontes, 2006.

KOBASHI, N. I. **A Organização e a Transferência de Informação Documentária: problemas e perspectivas.** A organização do conhecimento e dos sistemas de classificação. Brasília: IBICT, 1996.

KOBASHI, Nair Yumiko. **Fundamentos semânticos e pragmáticos da construção de instrumentos de representação de informação.** Datagramazero: Revista de Ciência da Informação - v.8 n.6, p. 1-10, 2007.

LAKATOS, Eva Maria, MARCONI, Marina de Andrade. **Metodologia científica.** 4. ed. São Paulo: Atlas, 2001.

LALANDE, A. **Vocabulário Técnico e Crítico da Filosofia.** São Paulo: Martins Fontes, 1993. 1336p.

LANGLOTZ, C. P. **Radlex: a new method for indexing on line education materials.** Radiographics. Vol. 26(6): 1595-1597, nov-dec, 2006.

LE GUERN, Michel. **Lês Descripteurs d'un système documentaire;** essai de definition. Actes du colloque traitement automatique des langues naturelles et systèmes documentaires, 1984. p. 162-169

LENAT, D. B.; GUHA, R.V. **Building large knowledge-based systems.** Massachussets: Addison-Wesley, 1990. 372 p.

LEVY, Pierre. **L'intelligence colletive: pour une antropogie du cyberpece.** Paris: La Découvert, 1997.

LIMA, J. C. de; CARVALHO, C. L. de. **Ontologias – OWL (Web Ontology Language).** Relatório Técnico, Univesidade Federal de Goiás, 2005.

MARTELETO, Regina Maria. **Conhecimento e socieade: pressupostos da antropologia da informação.** IN: AQUINO, Mirian de Albuquerque (Org.). O campo da ciência da Informação: gênese conexões e especialidades. João Pessoa: UFPb, 2002. 264p.

MENEZES, Paulo Blauth. **Linguagens Formais e Autômatos.** 5.ed. Porto Alegre: Bookman: Instituto de Informática da UFRGS, 2008.

MINAYO, Maria Cecília de Souza et al. (Org.) **Pesquisa social: teoria, método e criatividade.** Rio de Janeiro: Vozes, 1999.

MORA, José Ferrater. **Dicionário de Filosofia.** Bueno Aires: sudamericana, 1971.

MORRIS, Charles W. **Fundamentos da Teoria dos Signos.** São Paulo: USP, 1976.

MORTARI, Cezar A. **Introdução à Lógica.** São Paulo: UNESP, 2001.

MOURA, M. A. **Informação, Ferramentas Ontológicas e Redes Sociais: a interoperabilidade na construção de tesouros e Ontologias**. IN: IX Congresso ISKO-España, Valencia, março de 2009.

NEVES, J. Luis. Pesquisa Qualitativa – características, usos e possibilidades. **Caderno de Pesquisas em Administração**, São Paulo, v.1, n.3, 2º Sem./ 1996.

NÖTH, W. **Panorama da semiótica: de Platão a Peirce**. 4ªed. São Paulo: Annablume, 2009.

NOY, Natalya Fridman; FERGERSON, Ray W.; MUSEN, Mark A. **The knowledge model of Protégé-2000: combining interoperability and flexibility**. International Conference on Knowledge Engineering and Knowledge Management (EKAW'2000), Juan-les-Pins, France. 2000. Disponível em: <http://eprints.kfupm.edu.sa/70352/1/70352.pdf> . Acesso em: 15.Jun.2010.

OLIVEIRA, Silvio Luiz de. **Tratado de Metodologia Científica**. São Paulo: Pioneira, 2002.

PAIVA, Daniel C. de. **Modelagem e Simulação de Multidões Humanas em Situações da Vida Cotidiana usando Ontologias**. 145f. Dissertação (Mestrado em Computação Aplicada) - UNISINOS, São Leopoldo – RS, 2006.

PEIRCE, Charles S. **Semiótica**. São Paulo: Perspectiva, 1976.

PEIRCE, Charles S. **Semiótica**. São Paulo: Perspectiva, 1977.

PICKLER, M. E. V. . **Web Semântica: Ontologias como ferramentas de representação do conhecimento**. Perspectivas em Ciência da Informação, v. 12, p. 65-83, 2007

RIBEIRO, Adagenor Lobato. **As camadas da Arquitetura da Web Semântica** (2008). Disponível em: <http://adagenor.blogspot.com/2008/03/as-camadas-da-arquitetura-da-web.html>. Acesso em: 10. maio. 2010.

RIBEIRO, Guilherme A. S. **Pet Comp: Web Semântica**. Disponível em: <http://pet.ufma.br>. Acesso em 12.maio.2012.

RUBIN, Daniel L.; NOY, Natalya F.; MUSEN, Mark A. Protégé: A Tool for Managing and Using Terminology in Radiology. Applications. **Journal of Digital Imaging**, Vol. 0, N. 0 (Month), 2007: pp 1-13.

SALES, L. F. **Ontologias de domínio: um estudo das relações conceituais e sua aplicação**. 142 f. Dissertação (Mestrado) – Universidade Federal Fluminense, Niterói, 2006.

SALMON, Wesley C. **Lógica**. Rio de Janeiro: LCT, 2002.

SANTAELLA, L. **A Teoria Geral dos Signos**: semiose e autogeração. São Paulo: Ática, 1995

SANTAELLA, L. **Matrizes da Linguagem e Pensamento**: sonora, visual e verbal. 3 ed. São Paul: Iluminuras: Fapesp, 2005.

SANTAELLA, L. **O Método Anticarteziano de C. S. Peirce**. São Paulo: Unesp, 2004.

SANTAELLA, L. **O que é semiótica?** São Paulo: Brasiliense, 2003.

SANTAELLA, L. **Roteiro para a Leitura de Peirce**. v.2. O Falar da Linguagem, Série Linguagem. São Paulo: Lovise, 1997. p. 93-114.

SANTAELLA, L. **Semiótica Aplicada**. São Paulo: Thomson, 2002.

SANTOS, M.; FUJINO, A. **Serviços de Anotação em Imagens Médicas Baseado em Linguagem de Marcação**. In: Anais do XI ENANCIB, 2010, Rio de Janeiro-RJ.

SAUSSURE F. de. **Curso de Lingüística Geral**. São Paulo: Cultrix, 2006.

SEMPREBOM, Thiago; CAMADA; Marcos Y. ; MENDONÇA, Igor. **Ontologias e Protégé**. UFSC, Programa de Pós-Graduação em Engenharia de Automação e Sistemas, 2007. Disponível em: <http://www.das.ufsc.br/~gb/pg-ia/Protege07/Ontologias.pdf> . Acesso em: 07. Jul. 2010.

SILVA, D. L.; SOUZA, R.R.; ALMEIDA, M.B. Ontologias e Vocabulários Controlados: comparação de metodologias para construção. **Rev. Cin. Inf.**, Brasília v.37, n.3, p.60-75, set/dez, 2008.

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

SOCIEDADE PAULISTA DE RADIOLOGIA E DIAGNOSTICO POR IMAGEM. Disponível em <http://www.spr.org.br> . Acesso em: junho de 2012.

SOUZA, R. A.; HINTZE, A. C. J. **Pragmatismo e Lingüística: interfaces e intersecções**. Revista Eletrônica de Filosofia, São Paulo, v.7, n.2, julho-dez., 2010, p. 108-120.

SOUZA, R. Rocha; ALVARENGA, L.. A Web Semântica e suas contribuições para a ciência da informação. **Ci. Inf., Brasília**, v. 33, n. 1, p. 132-141, jan./abr. 2004.

SOWA, John F. **Building, sharing and merging ontologies**. [S. 1.: s. n.], 1999. Disponível em: <http://www.jfsowa.com/ontology/ontosar.htm> . Acessado em: 24. maio. 2012.

SOWA, John F. **Knowledge Representation: Logical, Philosophical, and Computational Foundations**. London: Brooks/Cole, 2000.

SOWA, John F. **Concepts in the Lexicon: Introduction**. Disponível em <http://users.bestweb.net/~sowa/ontology/lexicon.html>. Acessado em nov. de 2011.

STANFORD. **Guia do protégé-2000**. Disponível em http://protege.stanford.edu/doc/users_guide/index.html. Acessado em: 02. maio. 2010.

TUGENDHAT, Ernst. **Propedêutica lógico-semântica**. Petrópolis: Vozes, 1996.

WACHE, H. *et al.* Ontology-based integration of information – a survey of existing approaches. *In*: IJCAI, **WORKSHOP ON ONTOLOGIES AND INFORMATION SHARING**, 2001. [S. l. : s. n.] Disponível em: <http://citeseer.nj.nec.com/wache01ontologybased.html> . Acesso em: 01 maio 2002.

Apêndice

Apêndice 1: Elementos de composição da Ferramenta **Protégé 2000**

Como se pode observar, a interatividade e a flexibilidade encontradas no Protégé decorrem de uma série de subsistemas que se integram para garantir a obtenção de resultados eficazes. Segundo Noy (2000), o Protégé-2000 é um programa de organização de conhecimento baseado em frames, definidos pela autora como sendo “os blocos principais de uma base de conhecimentos”. Através desses frames a Ontologia oriunda do Protégé consiste basicamente de classes, slots e instâncias.

O campo em destaque que aparece quando se inicia um projeto é o campo "Classes" (etiqueta **Classes Tab**), considerado o mais importante. As Classes são tratadas no sistema como conceitos concretos do domínio e vão corresponder aos objetos, ou tipos de objetos do domínio. Os Atributos dessas classes são descritos por *slots* (etiqueta **Slots Tab**). As instâncias (etiqueta **Instances Tab**) são os dados inseridos na base de conhecimento.

A inserção de dados no Protégé é feita por meio de um formulário padrão, utilizado, dentre outras coisas, para inserir as instâncias. Os formulários são constituídos de um campo de entrada de dados para cada *Slot* relacionado a uma Classe, denominado de *widget* (forma gráfica). O formulário possui diversos tipos de formas gráficas, para facilitar a manipulação de vários tipos de valores de *Slots*. No entanto, o programa flexibiliza o uso dos formulários. Podem ser criados diferentes formulários, de acordo com as necessidades de cada desenvolvedor de ontologia. Os campos estão dispostos na interface do sistema em estrutura de “abas”, que denominaremos aqui de etiquetas, que facilitam o manuseio da ferramenta durante o processo.

Campo CLASSES

Ao acionar esse campo, visualiza-se uma árvore hierárquica com os seguintes itens de classe:

- 1- **Superclasse:** se encontra no topo da hierarquia da árvore taxonômica e corresponde aos conceitos mais gerais do domínio trabalhado na Ontologia;
- 2- **Classe:** são todos os conceitos do domínio do discurso da Ontologia.
- 3- **Subclasse:** são os conceitos mais específicos ou relacionados aos conceitos de uma classe
- 4- **Metaclasse:** modelo de classes que são as suas instâncias. Descreve como uma classe que instancia esse modelo vai olhar, por exemplo, os slots que ela própria terá e quais são as restrições para os valores desses slots. São meta-informações de uma classe.

Noy (2000) explica que as classes no Protégé-2000 “constituem uma hierarquia taxonômica”. Ou seja, se uma classe denominada X é uma subclasse de uma classe denominada Y, então, cada instância de X é também uma instância de Y. A autora afirma que o Protégé suporta herança múltipla de classes, quando uma classe pode ter mais de uma superclasse. Um exemplo para ilustrar melhor o contexto hierárquico do sistema: uma classe denominada **Editor**, que representa os editores de jornais, é simultaneamente uma subclasse da classe denominada **Empregado** e da classe denominada **Autor**, uma vez que um editor de jornal é tanto um empregado como um dos autores do jornal.

A figura abaixo mostra a disposição dos elementos do item classe da Ontologia Protégé, que se apresentam na tela em estrutura de árvore hierárquica.

Figura 1: Disposição dos elementos do item classe do Protégé
 Fonte: Semprebom, ET ali, 2007.

Pode-se perceber que ao acionar a etiqueta Classes automaticamente aparece uma janela que auxilia na visualização, na edição e na criação das classes de um projeto de Ontologia, conforme figura abaixo.

Figura 2: Janela de visualização das classes Protégé

Fonte: http://protege.stanford.edu/doc/users_guide

Nessa janela é possível visualizar a estrutura dos termos ou classes criadas para a base de conhecimento em construção e suas respectivas relações, como uma árvore hierárquica, mostrando as classes e as subclasses.

Todas as classes são mostradas como descendentes, direta ou indiretamente, da superclasse do sistema, denominada **Thing** (coisa). Os ícones ao lado de cada classe, se acionados, mostram informações adicionais sobre a classe em destaque. Cada ícone à esquerda ou direita do nome da classe tem função específica:

Quadro 1: Ícones e suas respectivas descrições

Indica que todas as subclasses diretas da classe são mostradas no momento. Clicando no ícone, podem-se esconder as subclasses.
Indica que a classe tem subclasses diretas que não são exibidas no momento. Clicando no ícone podem-se exibir as subclasses.
Nenhum ícone	A ausência de um ou indica ausência de subclasses na classe selecionada.
A cor amarela do ícone indica que a classe pode ser editada.
A cor verde do ícone indica que a classe é uma metaclassse.
A cor pálida do ícone indica que a classe foi incluída a partir de um outro projeto, o que indica que a classe não pode ser editada.
Uma classe pode ter um dos dois papéis: concreto ou abstrato. Classes concretas podem ter instâncias diretas; classes abstract não pode ter instâncias diretas. A ausência do ícone significa que a classe é concreta .
A classe tem múltiplas superclasses. O Protégé-2000 permite que as classes possam ter mais de uma superclasse. O ícone significa que a classe tem múltiplas superclasses. A ausência do ícone significa que a classe tem apenas uma superclasse.
Indica que a classe está oculta. A ausência do ícone significa que a classe não é escondida. Pode-se escolher se se quer ou não que as classes ocultas sejam exibidas no painel de relações de classe.

Fonte: http://protege.stanford.edu/doc/users_guide

A exibição dos ícones pode ser controlada, a partir da configuração no menu do programa, alternando para opções mais apropriadas.

Além dos ícones, encontram-se também os botões **VC** utilizados para a criação e/ou edição das classes. Localizados no canto superior direito da janela, os botões seguem todos os painéis do Protégé-2000 e têm as seguintes funções (STANFORD, 2010):

V (IEW): abrir o formulário de classe para a classe em destaque. Isto permite visualizar ou editar a mesma.

C (reate): criar uma nova classe como uma subclasse da classe destacada. .

🔍 **Voltar referências botão:** ver todos os objetos que fazem referência à classe em destaque.

✖ **Botão Excluir:** apagar a classe em destaque e todas as suas subclasses. .

O Protégé oferece um esquema de formulários de inserção de dados para a composição da Ontologia. Quando se deseja adicionar uma nova classe ou novos atributos de classes já inseridas na Ontologia, um formulário é aberto automaticamente para que tais ações sejam de fácil operação, conforme mostra a figura abaixo.

Name	Type	Cardinality	Other Facets
responsible_for	Instance	multiple	classes={Employee}
sections	Instance	multiple	classes={Section}
byname	String	single	
salary	Float	single	
date_hired	String	single	
current_job_title	String	single	
other_information	String	single	
name	String	single	
phone_number	String	single	

Figura 3: Janela de formulário do Protégé

Fonte: http://protege.stanford.edu/doc/users_guide

Os campos do formulário permitem: nomear uma classe; determinar a devida função da classe, especificando se é uma classe concreta ou abstrata, no campo **Role**; restringir os atributos de classe no campo **constraints**; inserir um texto descritivo para cada classe no campo de **documentação**; visualizar todos os slots herdados pela classe em construção, no painel **template slots**. O formulário possibilita ainda a inserção e/ou a exclusão de notas para a classe

trabalhada, por meio dos ícones , localizados no canto superior direito do formulário.

Campo SLOTS

Os *Slots* no Protégé-2000 têm a função de descrever todas as propriedades pertinentes às classes e instâncias, como o conteúdo de um livro ou o nome de um autor (NOY, 2000). Nas primeiras versões do Protégé, os *slots* eram utilizados somente dentro de classes. No entanto, com o desenvolvimento do programa, ele passou a ser um elemento independente, elevado ao mesmo nível de uma classe, para facilitar a inserção de informação no sistema. Assim, um *slot* passou a ser objeto de primeira classe, sendo definido independentemente de qualquer classe.

No Protégé, o painel dos *Slots* oferece uma janela cuja função permite criar, editar ou apenas visualizá-los. A janela é composta de três campos, conforme figura abaixo:

Figura 4: Janela dos Slots Protégé

Fonte: http://protege.stanford.edu/doc/users_guide

- a) Painel de Slots: no canto superior esquerdo da tela, permite a visualização de todos os slots criados na Ontologia, bem como a edição de cada item visualizado, a transferência entre slots e a exclusão de algum que não se adaptou ao projeto.
- b) Painel de Classes: no canto inferior esquerdo da tela, não é um campo de manipulação, mas uma janela apenas de exibição, cuja função é mostrar todas as classes que possuem um slot anexado.
- c) Painel do Formulário de slots: localizado em toda parte direita da tela, se configura como o painel de edição de slots, permitindo que se nomeie um slot, escolha sua cardinalidade e o tipo de valor, que se definam as restrições, os padrões inserindo uma breve descrição de cada um.

Obs.: O campo de cardinalidade permite especificar o número de valores permitidos e necessários para o *slot*. A configuração padrão permite que o *slot* possa ter no máximo um valor, ou seja, ele pode ter um valor ou nenhum valor (STANFORD, 2010).

Cardinality

☒ required at least

☒ multiple at most

Os ícones encontrados ao lado dos termos e os botões **VC** seguem as mesmas especificações do Campo Classes, descritos anteriormente. Eles auxiliam na edição e na organização das informações inseridas.

Campo FORMULÁRIOS

Assim como as Classes, as Instâncias se apresentam num projeto de Ontologia Protégé como um importante elemento para os especialistas do domínio pois é através delas que se atribuem os devidos valores conceituais às

Classes criadas. Isso é possível pelo preenchimento dos formulários-padrão que o Protégé-2000 gera para cada Classe. Acionando a etiqueta **Forms**, conforme figura abaixo, pode-se personalizar a forma como o domínio será visto e tratado.

Figura 5: Janela Formulários-padrão Protégé

Fonte: http://protege.stanford.edu/doc/users_guide

Assim, as formas conceituais que forem sendo criadas para as classes, no campo **Forms**, serão utilizadas também para introduzir as instâncias no campo Instâncias. Nesse campo também é possível criar várias interfaces de usuário para a mesma Ontologia - as formas de aquisição de conhecimento podem ser personalizadas para grupos de usuários (STANFORD, 2010).

Ao abrir o campo **Forms** aparece uma única janela na qual se pode visualizar e editar as formas apropriadas de protótipo. Essas formas, desenhadas e criadas a partir das necessidades elencadas do projeto de Ontologia, poderão ser visualizadas em seu formato final, ou seja, no campo das Instâncias.

A janela de interface do campo **Forms** contém apenas dois painéis (STANFORD, 2010):

1. O [painel de Formulários](#) localizado à esquerda do campo, que mostra a hierarquia de todas as classes, permite limpar todas as personalizações do formulário, bem como criar um formulário com um dos seguintes layouts predefinidos: o esquema padrão, o layout de um formulário que tenha sido criado por um pai, ou o layout de um formulário para outra classe.

2. O [formulário de edição](#) mostra o layout do formulário associado à classe selecionada. Cada slot na classe está associada a uma interface widget-usuário no formulário. Se ainda não foi criada ou modificada a forma, o Protégé-2000 usa um layout padrão com base no tipo de slot e cardinalidade.

Figura 6: Janela do formulário associado à classe selecionada.

Fonte: http://protege.stanford.edu/doc/users_guide

Campo INSTÂNCIAS

Nesse campo encontra-se uma janela na qual é possível ver, criar e editar **Instâncias**, ou seja, conceitos de classes do domínio trabalhado, slots propriedades de classes e as possíveis relações entre eles, bem como os dados reais do modelo de instâncias.

Ao acionar a etiqueta **Instances** uma única janela se abre, a qual é composta de três painéis, conforme mostra figura abaixo:

Figura 7: Janela das Instâncias Protégé

Fonte: <http://mimwiki.med.up.pt/>

1. O **painel das Classes**: localizado no canto superior direito da janela, é possível visualizar as classes criadas. No entanto, esse campo é somente para visualização, não sendo permitida a edição ou reorganização das mesmas. Para tal, é necessário ir até o campo propriamente dito das Classes.
2. O **painel Direct Instances**: localizado no centro da janela, mostra todas as instâncias diretas, se houver, para a classe selecionada. Além da visualização, é permitido editar, criar e excluir instâncias diretas.

Uma Instância direta é aquela que foi criada diretamente sob a classe. Contudo, por causa da herança, uma instância direta de uma classe é também um exemplo de todas as superclasses da referida classe. Portanto, uma instância pode ser uma

instância de muitas classes, porém, só pode ser uma instância *direta* de uma única classe. (STANFORD, 2010)

Esse painel tem três componentes:

- Os botões **V** **C** **X** do campo **Instance**, permitem criar, editar e excluir instâncias.
- A janela **Instance**, que exibe a lista alfabética das instâncias para a classe selecionada.
- O botão de pesquisa, localizado no canto inferior direito da janela, permite localizar uma instância na janela Instâncias, digitando-se o nome da instância e clicando no binóculo.

Figura 8: Canto inferior direito da janela Instancias

Fonte: http://protege.stanford.edu/doc/users_guide

3. O **painel formulário das Instâncias**: localizado no canto direito da janela, é visualizado toda vez que uma instância é selecionada. O painel de edição contém o formulário para a instância selecionada, que exibe todos os slots aplicáveis ao caso e permite editá-los.

Anexos

Anexo 1: Coleta de dados da Classe Imaging Observation Characteristic

Taxonomia da classe Imaging Observation Characteristic

IMAGING OBSERVATION CHARACTERISTIC

-
- aggressiveness modifier
 - aggressive
 - benign
 - expansile
 - high grade
 - higher probability of malignancy
 - indolent
 - invasive
 - low grade
 - malignant
 - noninvasive
 - typically benign
- composition modifier
 - calcified
 - cartilaginous
 - caseating
 - cavitary
 - collagenous
 - complex
 - cystic
 - dessicated
 - extremely dense
 - fat containing
 - gas containing
 - gaseous
 - hemorrhagic
 - heterogeneously dense
 - homogeneous
 - lipid containing
 - liquefied
 - mucinous
 - non-caseating
 - non-hemorrhagic
 - nuclear
 - osseous
 - ossified
 - predominantly fatty
 - proteinaceous
 - RID6060
 - scattered fibroglandular densities
 - semisolid
 - serous

- [simple](#)
 - [solid](#)
- [disease origin modifier](#)
 - [congenital](#)
 - [dystrophic](#)
 - [iatrogenic](#)
 - [idiopathic](#)
 - [non-syndromic](#)
 - [secretory](#)
 - [spontaneous](#)
 - [syndromic](#)
- [distribution pattern](#)
 - [coalescent](#)
 - [diffuse](#)
 - [focal](#)
 - [grouped](#)
 - [linear](#)
 - [multifocal](#)
 - [patchy](#)
 - [regional](#)
 - [scattered](#)
- [imaging observation location modifier](#)
 - [extramural](#)
 - [extraperitoneal](#)
 - [intramural](#)
 - [intraperitoneal](#)
 - [myocardium layer](#)
 - [RID6251](#)
- [modality-related characteristic](#)
 - [attenuation characteristic](#)
 - [density characteristic](#)
 - [echogenicity characteristic](#)
 - [enhancement characteristic](#)
 - [signal characteristic](#)
- [morphologic characteristic](#)
 - [aerated](#)
 - [amorphous](#)
 - [angular](#)
 - [aphthous](#)
 - [asymmetrical](#)
 - [atrophic](#)
 - [beaded](#)
 - [bilobed](#)
 - [blunted](#)
 - [bulky](#)
 - [coarse](#)
 - [coarse heterogeneous](#)
 - [compressed](#)
 - [conglomerate](#)
 - [consolidated](#)

- [curved](#)
- [discoid](#)
- [distorted](#)
- [eggshell](#)
- [emphysematous](#)
- [eroded](#)
- [fatty](#)
- [fibrous](#)
- [fine linear](#)
- [fine pleomorphic](#)
- [flat](#)
- [flattened](#)
- [fungating](#)
- [fusiform](#)
- [gently lobulated](#)
- [globular](#)
- [hydrated](#)
- [hyperinflated](#)
- [impinged](#)
- [indented](#)
- [invaginated](#)
- [irregular](#)
- [lacunar](#)
- [large rod-like](#)
- [lipid](#)
- [lobular](#)
- [long](#)
- [lucent-centered](#)
- [mixed](#)
- [multilocular](#)
- [narrowed](#)
- [nodular](#)
- [obstructed](#)
- [occluded](#)
- [oval](#)
- [patent](#)
- [patulous](#)
- [pedunculated](#)
- [plaquelike](#)
- [plate-like](#)
- [plateau](#)
- [pleomorphic](#)
- [polygonal](#)
- [polyp morphology characteristic](#)
- [polypoid](#)
- [puckered](#)
- [punctate](#)
- [reconstructed](#)
- [reticular](#)
- [retractile](#)

- [round](#)
- [ruptured](#)
- [scarred](#)
- [sclerosing](#)
- [septated](#)
- [serpentine](#)
- [serpiginous](#)
- [shallow](#)
- [sharp](#)
- [short](#)
- [smooth](#)
- [spoke-wheel](#)
- [straightened](#)
- [symmetrical](#)
- [tethered](#)
- [thickened](#)
- [thinned](#)
- [tortuous](#)
- [truncated](#)
- [tubular](#)
- [ulcerated](#)
- [unilocular](#)
- [varicoid](#)
- [wedge-shaped](#)
- [motion characteristic](#)
 - [immobile](#)
 - [in transit](#)
 - [mobile](#)
 - [pulsatile](#)
 - [sliding](#)
- [obscured](#)
- [prominent](#)
- [radiopacity](#)
 - [radiolucent](#)
 - [radiopaque](#)
- [temporal modifier](#)
 - [acute](#)
 - [chronic](#)
 - [fixed](#)
 - [new](#)
 - [old](#)
 - [onset modifier](#)
 - [permanent](#)
 - [persistent](#)
 - [postoperative](#)
 - [preoperative](#)
 - [previous](#)
 - [repetitive](#)
 - [residual](#)
 - [reversible](#)

- [temporary](#)
- [trend modifier](#)

Detalhes de alguns termos selecionados

01 Preferred Name	imaging observation characteristic
Synonyms (Synonym)	Bildbeobachtungseigenschaft
ID	RID7464
Full Id	http://purl.bioontology.org/ontology/RID/RID7464
Direct Superclasses	Radlex descriptor
Direct Type	radlex metaclass
Name	RID7464
Role	Concrete
Preferred Name	imaging observation characteristic
Synonym	Bildbeobachtungseigenschaft
Direct Subclasses	temporal modifier aggressiveness modifier disease origin modifier distribution pattern imaging observation location modifier morphologic characteristic composition modifier motion characteristic obscured radiopacity prominent modality-related characteristic
Is A	Radlex descriptor

02 Preferred Name (Preferred_name)	aggressiveness modifier
Synonyms (Synonym)	Modifikator Aggressivität
ID	RID5675
Full Id	http://purl.bioontology.org/ontology/RID/RID5675
Direct Superclasses	imaging observation characteristic
Direct Type	radlex metaclass
Name	RID5675
Role	Concrete
Preferred Name	aggressiveness modifier
Synonym	Modifikator Aggressivität
Direct Subclasses	expansile higher probability of malignancy malignant

	noninvasive benign invasive high grade aggressive low grade indolent typically benign
Is A	imaging observation characteristic

03 Preferred Name	aggressive
Synonyms (Synonym)	aggressiv
ID	RID5678
Full Id	http://purl.bioontology.org/ontology/RID/RID5678
Direct Superclasses	aggressiveness modifier
Direct Type	radlex_metaclass
Name	RID5678
Role	Concrete
Preferred Name	aggressive
Synonym	aggressiv
Is A	aggressiveness modifier

04 Preferred Name	benign
ID	RID15654
Full Id	http://purl.bioontology.org/ontology/RID/RID15654
Direct Superclasses	aggressiveness modifier
Direct Type	radlex_metaclass
Name	RID15654
Role	Concrete
Preferred Name	benign
Is A	aggressiveness modifier

05 Preferred Name	expansile
Synonyms (Synonym)	expansil
ID	RID5682
Full Id	http://purl.bioontology.org/ontology/RID/RID5682
Direct Superclasses	aggressiveness modifier
Direct Type	radlex_metaclass
Name	RID5682
Role	Concrete
Preferred Name	expansile

Synonym	expansil
Is A	aggressiveness modifier

OBS.: Anexo completo no CD-ROM

Anexo 2: Coleta de dados da Classe Pathophysiologic

Taxonomia da classe Pathophysiologic Process Modifier

- [pathophysiologic process modifier](#)
 - [aortic dissection modifier](#)
 - [arthritis modifier](#)
 - [amount of joint involvement](#)
 - [monoarticular](#)
 - [pauciarticular](#)
 - [polyarticular](#)
 - [arthritis serology](#)
 - [seronegative](#)
 - [seropositive](#)
 - [periarticular](#)
 - [bone pathophysiologic process modifier](#)
 - [bone forming characteristic](#)
 - [osteolytic](#)
 - [osteosclerotic](#)
 - [cardiomyopathy descriptor](#)
 - [congestive](#)
 - [hypertrophic](#)
 - [infiltrative](#)
 - [ischemic](#)
 - [restrictive](#)
 - [effusion modifier](#)
 - [chylous](#)
 - [exudate](#)
 - [head pathophysiologic process observation modifier](#)
 - [alobar](#)
 - [gyriform](#)
 - [hemispheric](#)
 - [leptomeningeal](#)
 - [pachymeningeal](#)
 - [semilobar](#)
 - [suprasellar](#)
 - [territorial](#)
 - [transependymal](#)
 - [watershed](#)
 - [hormonal pathophysiologic process modifier](#)
 - [feminizing](#)
 - [virilizing](#)
 - [androgen producing](#)
 - [mass modifier](#)
 - [exophytic](#)
 - [intraparenchymal](#)

- [renal ectopia modifier](#)
 - [crossed](#)
 - [fused](#)
 - [uncrossed](#)
 - [unfused](#)
- [varices modifier](#)
 - [downhill](#)
 - [uphill](#)

Detalhes de alguns termos selecionados

1 Preferred Name (Preferred_name)	Pathophysiologic process modifier
Synonyms (Synonym)	Pathophysiologischer Prozess Beobachtungsmodifikator
ID	RID7466
Full Id	http://purl.bioontology.org/ontology/RID/RID7466
Direct Superclasses	Radlex descriptor
Direct Type	radlex metaclass
Name	RID7466
Role	Concrete
Preferred Name	pathophysiologic process modifier
Synonym	Pathophysiologischer Prozess Beobachtungsmodifikator
Direct Subclasses	mass modifier cardiomyopathy descriptor varices modifier head pathophysiologic process observation modifier hormonal pathophysiologic process modifier bone pathophysiologic process modifier effusion modifier arthritis modifier aortic dissection modifier renal ectopia modifier
Is A	Radlex descriptor

2 Preferred Name (Preferred_name)	aortic dissection modifier
Synonyms (Synonym)	Modifikator Aortendisektion
ID	RID6369
Full Id	http://purl.bioontology.org/ontology/RID/RID6369
Direct Superclasses	pathophysiologic process modifier
Direct Type	radlex metaclass
Name	RID6369
Role	Concrete

Preferred Name	aortic dissection modifier
Synonym	Modifikator Aortendissektion
Direct Subclasses	Stanford type A DeBakey type I Stanford type B DeBakey type II DeBakey type III
Is A	pathophysiologic process modifier

3 Preferred Name (Preferred_name)	DeBakey type I
Synonyms (Synonym)	type I (DeBakey)
ID	RID6372
Full Id	http://purl.bioontology.org/ontology/RID/RID6372
Direct Superclasses	aortic dissection modifier
Direct Type	radlex_metaclass
Name	RID6372
Role	Concrete
Preferred Name	DeBakey type I
Synonym	type I (DeBakey)
Is A	aortic dissection modifier

4 Preferred Name (Preferred_name)	DeBakey type II
Synonyms (Synonym)	type II (DeBakey)
ID	RID6373
Full Id	http://purl.bioontology.org/ontology/RID/RID6373
Direct Superclasses	aortic dissection modifier
Direct Type	radlex_metaclass
Name	RID6373
Role	Concrete
Preferred Name	DeBakey type II
Synonym	type II (DeBakey)
Is A	aortic dissection modifier

OBS.: Anexo completo no CD-ROM

Anexo 3: Coleta de dados da Classe Anatomical Cluster

Taxonomia da classe Anatomical Cluster

- **Anatomical Cluster**
 - [anatomical compartment](#)
 - [compartment of trunk](#)
 - [body compartment subdivision](#)
 - [compartment of abdomen](#)
 - [subdivision of abdominal compartment](#)
 - [compartment of female pelvis](#)
 - [compartment of male pelvis](#)
 - [left deep perineal compartment](#)
 - [right deep perineal compartment](#)
 - [left extraperitoneal compartment](#)
 - [right extraperitoneal compartment](#)
 - [left retropubic compartment](#)
 - [right retropubic compartment](#)
 - [left superficial perineal compartment](#)
 - [right superficial perineal compartment](#)
 - [subdivision of thoracic compartment](#)
- [anatomical conduit](#)
 - [greater sciatic foramen](#)
 - [hypoglossal canal](#)
 - [infrapiriform foramen](#)
 - [inguinal canal](#)
 - [lesser sciatic foramen](#)
 - [pelvic inlet](#)
 - [pelvic outlet](#)
 - [suprapiriform foramen](#)
 - [thoracic outlet](#)
- [anatomical junction](#)
 - [decussation](#)
 - [gastroesophageal junction](#)
 - [joint](#)
 - [labial commissure of mouth](#)
 - [lumbosacral junction](#)
 - [nerve plexus](#)
 - [thoracolumbar junction](#)
 - [ureteropelvic junction](#)
 - [ureterovesical junction](#)
 - [zone of continuity](#)
- [body part subdivision cluster](#)
 - [pudendum](#)
 -

- [bony palate](#)
- [carpal bridge region](#)
- [cell part cluster](#)
 - [cell part cluster of neuraxis](#)
 - [commissural white matter tract](#)
 - [long association fiber](#)
 - [short association fiber](#)
- [compartment subdivision](#)
 - [subdivision of epidural space](#)
- [gestational structure](#)
 - [embryonic structure](#)
 - [extra-embryonic structure](#)
- [heterogeneous cluster](#)
 - [content of abdomen](#)
 - [content of abdomen proper](#)
 - [content of pelvis](#)
 - [larynx](#)
 - [left side of pelvic wall](#)
 - [pelvic brim](#)
 - [pelvic wall](#)
 - [pharynx](#)
 - [posterior abdominal wall](#)
 - [right side of pelvic wall](#)
 - [subdivision of larynx](#)
 - [subdivision of pharynx](#)
 - [urogenital diaphragm](#)
 - [vestibulo-auditory system](#)
 - [vestibulocochlear organ](#)
 - [visual system](#)
 - [wall of gastrointestinal tract](#)
 - [wall of lower gastrointestinal tract](#)
 - [wall of upper gastrointestinal tract](#)
- [inner ear](#)
- [meningeal cluster](#)
 - [dura mater](#)
- [middle ear](#)
 - [middle ear cavity](#)
- [organ cluster](#)
 - [iliopsoas](#)
 - [pericardium](#)
- [organ part cluster](#)
 - [organ component cluster](#)
 - [organ subdivision cluster](#)
 - [wall of intestine](#)
- [palate](#)
- [pelvic floor](#)
- [renal collecting system](#)
 - [calyx of renal collecting system](#)
 - [infundibulum of renal collection system](#)
 - [renal pelvis](#)

- [scalp](#)
- [soft palate](#)
 - [uvula](#)
- [tongue](#)
- [upper esophageal sphincter](#)
- [vestigial embryonic structure](#)
 - [ligamentum teres of liver](#)
 - [urachus](#)

Detalhes de alguns termos selecionados

1 Preferred Name (Preferred_name)	anatomical cluster
Synonyms (Synonym)	anatomischer Block
ID	RID13290
Full Id	http://purl.bioontology.org/ontology/RID/RID13290
Direct Superclasses	anatomical structure
Direct Type	anatomy metaclass
Name	RID13290
Role	Concrete
Definition	Anatomical structure, which has as its parts a heterogeneous collection of organs, organ parts, cells, cell parts or body part subdivisions that are adjacent to, or continuous with one another; does not constitute a cell part, cell, tissue, organ, organ system or organ system subdivision, cardinal body part, body part subdivision or anatomical junction. Examples: joint, adnexa of uterus, root of lung, renal pedicle, back.
Fmaid	49443
Preferred Name	anatomical cluster
Synonym	anatomischer Block
Direct Subclasses	meningeal cluster heterogeneous cluster organ cluster organ part cluster anatomical conduit scalp carpal bridge region soft palate cell part cluster body part subdivision cluster bony palate vestigial embryonic structure anatomical junction pelvic floor

	inner ear upper esophageal sphincter renal collecting system compartment subdivision palate anatomical compartment middle ear tongue gestational structure
Is A	anatomical structure

2 Preferred Name (Preferred_name)	anatomical compartment
ID	RID29528
Full Id	http://purl.bioontology.org/ontology/RID/RID29528
Direct Superclasses	anatomical cluster
Direct Type	anatomy metaclass
Name	RID29528
Role	Concrete
Definition	Anatomical cluster which has as its parts one or more anatomical structures surrounding an anatomical compartment space and the contents of that space. Examples: mediastinal part of chest, anterior compartment of forearm, palm of hand(compartment), intercostal space, vertebral canal
Fmaid	9647
Preferred Name	anatomical compartment
Direct Subclasses	compartment of trunk
Is A	anatomical cluster

3 Preferred Name (Preferred_name)	compartment of trunk
Synonyms (Synonym)	trunk compartment
ID	RID29529
Full Id	http://purl.bioontology.org/ontology/RID/RID29529
Direct Superclasses	anatomical compartment
Direct Type	anatomy metaclass
Name	RID29529
Role	Concrete
Fmaid	58903
Preferred Name	compartment of trunk
Synonym	trunk compartment
Direct Subclasses	body compartment subdivision
Is A	anatomical compartment

4 Preferred Name (Preferred_name)	body compartment subdivision
ID	RID29530
Full Id	http://purl.bioontology.org/ontology/RID/RID29530
Direct Superclasses	compartment of trunk
Direct Type	anatomy metaclass
Name	RID29530
Role	Concrete
Definition	Compartment of trunk which is enclosed by a body wall subdivision and is demarcated from another body cavity subdivision by an anatomical structure or a conduit, it contains one or more serous sacs, viscera and other organs; together with the other body cavity subdivisions, it constitutes the body cavity. Examples: thoracic cavity, abdominopelvic cavity, abdominal cavity, pelvic cavity.
Fmaid	12236
Preferred Name	body compartment subdivision
Direct Subclasses	subdivision of thoracic compartment compartment of abdomen subdivision of abdominal compartment
Is A	compartment of trunk

OBS.: Anexo completo no CD-ROM

Anexo 4: Coleta de dados da Classe OBJECT

▪ Taxonomia da Classe Object

- [object](#)
 - [foreign body](#)
 - [bezoar](#)
 - [phytobezoar](#)
 - [trichobecaoar](#)
 - [contraceptive device](#)
 - [intrauterine device](#)
 - [decoration](#)
 - [body piercing](#)
 - [jewelry](#)
 - [tattoo](#)
 - [embolic material](#)
 - [adhesive](#)
 - [embolization coil](#)
 - [platinum embolization coil](#)
 - [steel embolization coil](#)
 - [particulate embolic agent](#)
 - [missile](#)
 - [birdshot](#)
 - [buckshot](#)
 - [bullet](#)
 - [shrapnel](#)
 - [personal item](#)
 - [coin](#)
 - [denture](#)
 - [knife](#)
 - [medical object](#)
 - [medical device](#)
 - [applicator](#)
 - [antenna applicator](#)
 - [bipolar applicator](#)
 - [cryoprobe applicator](#)
 - [electrode applicator](#)
 - [fiber applicator](#)
 - [internally cooled applicator](#)
 - [monopolar applicator](#)
 - [multi-tined applicator](#)
 - [perfusion applicator](#)
 - [straight cluster applicator](#)
 - [straight single applicator](#)
 - [balloon](#)
 - [compliant balloon](#)

- [cryoplasty balloon](#)
- [cutting balloon](#)
- [detachable balloon](#)
- [high-pressure balloon](#)
- [low-pressure balloon](#)
- [noncompliant balloon](#)
- [biopsy instrument](#)
 - [biopsy brush](#)
 - [core needle](#)
- [collimator](#)
 - [multihole collimator](#)
 - [pinhole collimator](#)
 - [xray collimator](#)
- [CT detector](#)
 - [solid-state ceramic scintillation detector](#)
 - [solid-state crystal scintillation detector](#)
 - [Xenon ionization chamber](#)
- [dilator](#)
 - [hydrophilic dilator](#)
 - [nonhydrophilic dilator](#)
- [drainage device](#)
- [filter](#)
 - [body filter](#)
 - [cardiac filter](#)
 - [head filter](#)
 - [k-space filter](#)
 - [pediatric filter](#)
- [imaging device](#)
 - [CT device](#)
 - [MR device](#)
 - [nuclear imaging device](#)
- [implantable device](#)
 - [battery pack](#)
 - [cardiac rhythm therapy device](#)
 - [cochlear implant](#)
 - [dental device](#)
 - [electrical stimulator](#)
 - [fiducial](#)
 - [hearing aid](#)
 - [implant](#)
 - [intracranial pressure monitor](#)
 - [ocular prosthesis](#)
 - [orbital prosthesis](#)
 - [orthopedic device](#)
 - [prosthesis](#)
 - [temporary pacemaker patch](#)

- [tube or catheter](#)
- [valve](#)
- [venous filter](#)
- [wire](#)
- [injector](#)
 - [gravity injector](#)
 - [power injector](#)
 - [pump injector](#)
- [interventional wire](#)
 - [hydrophilic wire](#)
 - [nonhydrophilic wire](#)
 - [nonstiff wire](#)
 - [stiff wire](#)
- [laser crystal](#)
- [magnetic resonance imaging coil](#)
 - [biopsy coil](#)
 - [flexible coil](#)
 - [gradient coil](#)
 - [internally inserted coil](#)
 - [microscopy coil](#)
 - [multinuclear coil](#)
 - [radiofrequency coil](#)
 - [shim coil](#)
- [nuclear camera](#)
 - [gamma camera](#)
 - [positron emission tomography camera](#)
- [nuclear diode](#)
 - [avalanche diode](#)
 - [PIN diode](#)
 - [silicon drift diode](#)
- [surgical equipment](#)
 - [Nanoknife](#)
- [surgical implement](#)
 - [bone graft substitute](#)
 - [ceramic surgical implement](#)
 - [forceps](#)
 - [free silicone](#)
 - [medication device](#)
 - [needle](#)
 - [sponge](#)
 - [surgical cement](#)
 - [surgical clip](#)
 - [surgical drain](#)
 - [surgical wire](#)
 - [temperature monitoring device](#)
- [transducer](#)
 - [ultrasound transducer](#)
- [portion of medical device](#)

- [laser tip](#)
 - [flexible diffuser laser tip](#)
 - [scattering dome laser tip](#)
- [needle sheath](#)
 - [curved sheath](#)
 - [hemostatic sheath](#)
 - [peel away sheath](#)
 - [straight sheath](#)
- [needle tip](#)
 - [bevel needle tip](#)
 - [diamond needle tip](#)
 - [serrated needle tip](#)
- [wire tip](#)
 - [angled wire tip](#)
 - [J wire tip](#)
 - [shapeable wire tip](#)
 - [steerable wire tip](#)
 - [straight wire tip](#)

Detalhes de alguns termos selecionados

1 Preferred Name (Preferred_name)	Object
ID	RID34861
Full Id	http://purl.bioontology.org/ontology/RID/RID34861
Direct Superclasses	RadLex entity
Direct Type	radlex_metaclass
Name	RID34861
Role	Concrete
Definition	A non-anatomical object entity
Preferred Name	Object
Direct Subclasses	medical object foreign body
Is A	RadLex entity

2 Preferred Name (Preferred_name)	foreign body
--	--------------

Synonyms (Synonym)	Fb foreign matter foreign object Fremdkörper
ID	RID5425
Full Id	http://purl.bioontology.org/ontology/RID/RID5425
Comment	child of finding
Direct Superclasses	object
Direct Type	radlex_metaclass
Name	RID5425
Role	Concrete
Definition	Objects that inadvertently enter the body or body cavities from the environment. [MeSH]
Member Of	set of BI-RADS ultrasound terms
Preferred Name	foreign body
Source	BI-RADS 4
Synonym	Fremdkörper foreign object foreign matter fb
Umls	C0016542
Umls Term	foreign body
Direct Subclasses	personal item missile embolic material decoration contraceptive device bezoar
Is A	object

3 Preferred Name (Preferred_name)	Bezoar
Synonyms (Synonym)	Bezoarstein
ID	RID5426
Full Id	http://purl.bioontology.org/ontology/RID/RID5426
Direct Superclasses	foreign body
Direct Type	radlex_metaclass
Name	RID5426
Role	Concrete
Definition	Concretions of swallowed hair, fruit or vegetable fibers, or similar substances found in the alimentary

	canal. [MeSH]
Preferred Name	Bezoar
Synonym	Bezoarstein
Umls	C0005332
Umls Term	Bezoar
Direct Subclasses	phytobezoar trichobecaoar
Is A	foreign body

4 Preferred Name (Preferred_name)	Phytobezoar
ID	RID5428
Full Id	http://purl.bioontology.org/ontology/RID/RID5428
Direct Superclasses	bezoar
Direct Type	radlex_metaclass
Name	RID5428
Role	Concrete
Preferred Name	Phytobezoar
Is A	bezoar

5 Preferred Name (Preferred_name)	Trichobecaoar
Synonyms (Synonym)	Trichobezoar
ID	RID5427
Full Id	http://purl.bioontology.org/ontology/RID/RID5427
Direct Superclasses	bezoar
Direct Type	radlex_metaclass
Name	RID5427
Role	Concrete
Preferred Name	Trichobecaoar
Synonym	Trichobezoar
Is A	bezoar

OBS.: Anexo completo no CD-ROM