

4 Ontologia

Nesse capítulo introduziremos a discussão sobre o que vem a ser uma ontologia dentro do domínio utilizado nesse trabalho. Veremos as principais vantagens do uso de ontologia, suas áreas de aplicação, suas classificações e metodologias de desenvolvimento.

4.1. O que é ontologia

O termo ontologia é originário da filosofia. Ontologia é um ramo da filosofia que lida com a natureza e a organização do ser. Esse termo foi introduzido por Aristóteles em *Metafísica*, IV, 1. No contexto da pesquisa em “ontologia”, filósofos tentam responder as questões “O que é um ser?” e “Quais são as características comuns de todos os seres?” (Maedche, 2002).

Esse termo foi recentemente adotado também pelas comunidades de inteligência artificial e gestão de conhecimento para se referir a conceitos e termos que podem ser usados para descrever alguma área do conhecimento ou construir uma representação desse.

A literatura sobre ontologias apresenta uma série de definições distintas. Essas diferentes definições apresentam pontos de vista distintos e até mesmo complementares para uma mesma realidade. Em Guarino (1997) é feita uma longa discussão sobre o significado preciso do termo dentro da Ciência da Computação, pois o seu significado preciso tende a variar conforme o objetivo do uso da ontologia.

Entre as várias definições de ontologia existentes, uma muito interessante é a apresentada por Fensel (2001) proveniente de T. R. Gruber em seu artigo “*A Translation Approach to Portable Ontology Specification*”:

“Uma ontologia é uma especificação formal explícita de uma conceitualização compartilhada.”

Nessa frase é importante explicitar-se o significado de algumas das palavras utilizadas. A palavra “conceitualização” refere-se a um modelo abstrato de algum fenômeno que identifique conceitos relevantes desse fenômeno. A palavra

“explícita” significa que os tipos de conceitos usados e as limitações do uso desses conceitos devem ser definidos de forma explícita. A palavra “formal” refere-se que a ontologia deve ser passível de ser processada por uma máquina. Por fim “compartilhada” reflete a noção de que a ontologia captura um conhecimento consensual, isto é, esse conhecimento não deve ser restrito a alguns indivíduos, mas aceito por um grupo de pessoas (Fensel, 2001).

Na descrição acima se pode constatar algumas características importantes que a ontologia deve possuir (a ontologia deve ser explícita, formal, e descrever um conhecimento comum a um grupo), porém não fala a respeito de como é a estrutura de uma ontologia.

Outra descrição do que vem a ser ontologia, que é de certa forma complementar a descrição acima, é a apresentada por Gómez-Pérez (1999) proveniente de Swartout e seu colegas presente no artigo “*Towards Distributed Use of Large-Scale Ontologies*”:

“Uma ontologia é um conjunto de termos ordenados hierarquicamente para descrever um domínio que pode ser usado como um esqueleto para uma base de conhecimentos.”

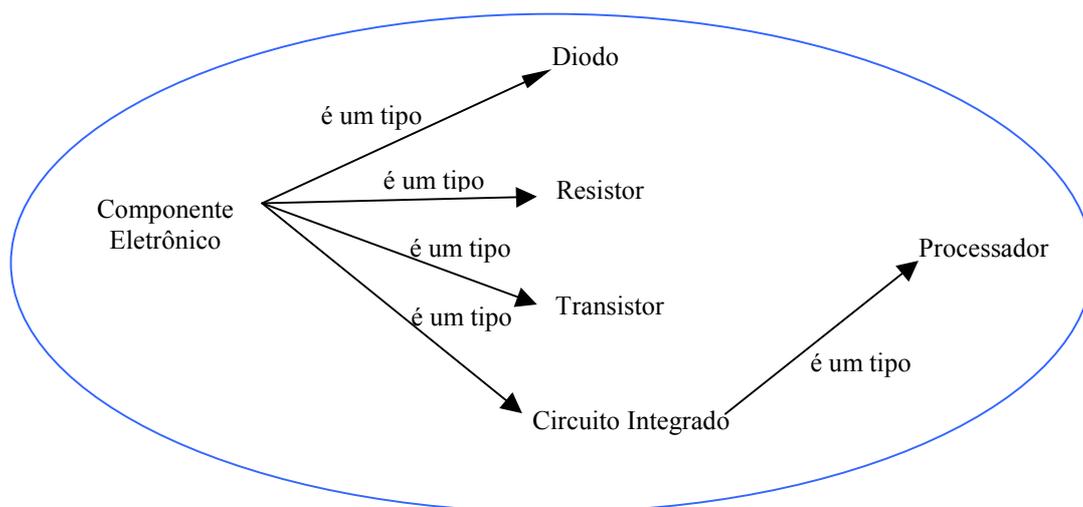
Na descrição de Swartout, já são dadas algumas informações a respeito da estrutura de uma ontologia, pois conforme essa descrição, uma ontologia deve possuir um conjunto de termos organizados com uma hierarquia associada, ou seja, uma taxinomia. Outro detalhe dessa descrição é apresentar uma das principais utilidades de uma ontologia que é a de servir como um “*schema*” para uma base de conhecimentos, visão essa muito comum dentro do ramo da gestão de conhecimentos.

Vale a pena ressaltar que a definição acima faz uma distinção importante sobre o que é uma ontologia e o que é uma base de conhecimentos. Uma ontologia provê uma estrutura básica na qual logo acima se pode construir uma base de conhecimentos. A ontologia fornece um conjunto de conceitos e termos para descrever um determinado domínio, enquanto a base de conhecimento usa esses termos para descrever uma determinada realidade. Caso essa realidade seja modificada a base de conhecimentos será modificada, porém a ontologia permanecerá inalterada, desde que o domínio se mantenha inalterado.

Para esclarecer a diferença entre ontologia e base de conhecimentos, a figura 12 apresenta um exemplo bem ilustrativo. Neste exemplo, pode-se definir uma ontologia para o domínio de circuitos eletrônicos. A ontologia resultante incluiria

conceitos específicos (por exemplo: circuito integrado, resistor, transistor, processador, diodo, etc.) e também relações entre esses conceitos (por exemplo: o processador é um tipo de circuito integrado, etc.). Na hora de colocar na base de conhecimentos um circuito de um rádio, ou seja, descrever a realidade de um rádio, esse pode ser formado por 10 transistores, 20 resistores, e 4 diodos. Já na hora de colocar na base de conhecimentos um circuito de uma calculadora, esse pode ser formado por 4 transistores, 3 resistores, e 5 diodos. Podemos observar que apesar da realidade descrita na base de conhecimento variar para o rádio e para a calculadora, a ontologia empregada continua sendo a mesma, ou seja, eu uso os mesmos conceitos para descrever as duas realidades.

Ontologia do domínio de circuitos eletrônicos



Base de conhecimento que utiliza a ontologia de circuitos eletrônicos

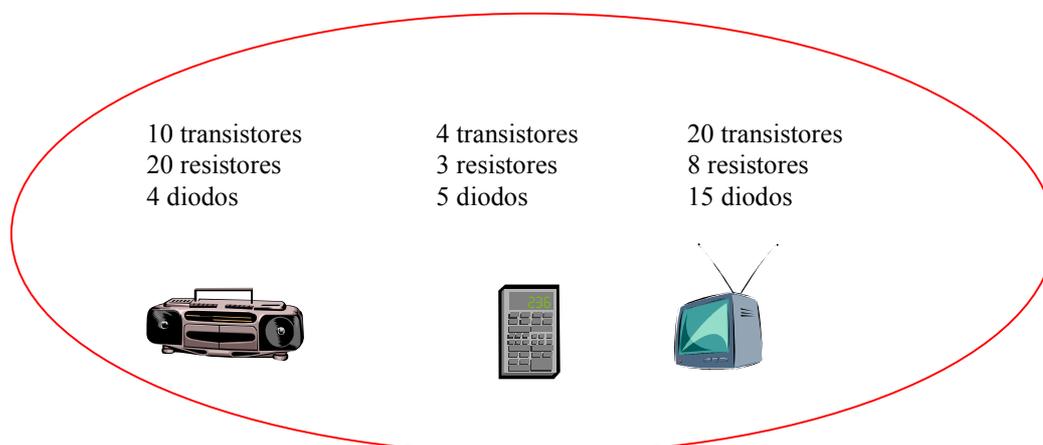


Figura 12 – Comparação entre ontologia e base de conhecimento.

Há outras definições a respeito do que vem a ser uma ontologia (Guarino, 1997), porém resumindo os principais pontos dessas definições pode-se dizer que uma ontologia é uma conceitualização compartilhada de um determinado domínio, conforme muito bem definido por T. R. Gruber. Ela é composta de um conjunto de conceitos dentro desse domínio, sendo esses organizados como uma taxinomia, e de relações entre esses conceitos. A ontologia possui também axiomas, ou seja, regras pertinentes ao domínio em questão.

4.2. **Vantagens do uso de ontologias**

Agora que já foi visto o que é uma ontologia, é importante que se destaque as vantagens de seu uso. A seguir é apresentada uma lista com as principais vantagens da utilização de ontologias na Ciência da Computação:

- Ontologias fornecem um vocabulário para representação do conhecimento. Esse vocabulário tem por trás uma conceitualização que o sustenta, evitando assim interpretações ambíguas desse vocabulário.
- Ontologias permitem o compartilhamento de conhecimento. Sendo assim, caso exista uma ontologia que modele adequadamente certo domínio de conhecimento, essa pode ser compartilhada e usada por pessoas que desenvolvam aplicações dentro desse domínio. Para exemplificar, considere que exista uma ontologia para o domínio de livrarias. Uma vez que essa ontologia está disponível, várias livrarias podem construir seus catálogos usando o vocabulário fornecido por essa ontologia sem a necessidade de refazer uma análise do domínio de livraria.
- Fornece uma descrição exata do conhecimento. Diferentemente da linguagem natural em que as palavras podem ter semântica totalmente diferente conforme o seu contexto, a ontologia por ser escrita em linguagem formal, não deixa espaço para o *gap* semântico existente na linguagem natural. Por exemplo, quando uma pessoa fala para outra a palavra “Globo” ela pode estar querendo falar a respeito de um corpo esférico, como também de um canal de televisão brasileiro. A interpretação da palavra pode ser atribuída a um conceito ou outro conforme o estado mental do indivíduo. Porém, se há uma conceitualização comum entre essas duas pessoas a possibilidade de mal entendido diminui muito. Por exemplo, se essas pessoas concordam em uma ontologia sobre o domínio de formas geométricas, possivelmente não haverá mal entendido.

- É possível fazer o mapeamento da linguagem da ontologia sem que com isso seja alterada a sua conceitualização, ou seja, uma mesma conceitualização pode ser expressa em várias línguas.
- Pode ser possível estender o uso de uma ontologia genérica de forma a que ela se adeque a um domínio específico. Por exemplo, se alguém precisa de uma ontologia sobre bicicletas para construir uma aplicação e só encontra uma ontologia sobre o domínio genérico de veículos, pode utilizar essa ontologia estendendo-a para o domínio específico da aplicação, que no caso são de bicicletas.

Essas são as principais vantagens da utilização de ontologias. Existem muitas outras vantagens, porém a maioria é decorrente dessas.

4.3. O que compõe uma ontologia

Entre as várias descrições sobre o que compõe uma ontologia (Maedche, 2002; Gómez-Pérez, 1999), a que melhor explicita os seus componentes é a proposta por Maedche (2002) principalmente devido ao seu alto grau de formalismo na definição da estrutura de uma ontologia. Apesar disso, a descrição dada por Gómez-Pérez (1999) é mais completa, por considerar que parte das instâncias é componente da ontologia. Sendo assim, para nós uma ontologia é formada pelos seguintes componentes:

- Um conjunto de conceitos e uma hierarquia entre esses conceitos, ou seja, uma taxinomia. Os conceitos podem ser abstratos (por ex.: força) ou concretos (por ex.: carro), elementares (por ex.: elétron) ou compostos (por ex.: átomo), reais ou fictícios. Um exemplo de taxinomia é o conceito *homem* ser um subconceito do conceito *pessoa*.
- Um conjunto de relacionamentos entre esses conceitos. Um exemplo de relacionamento entre os conceitos de *pessoa* e *carro* é o relacionamento de ser-dono.
- Um conjunto de funções. Uma função é um caso especial de relacionamento em que um conjunto de elementos tem uma relação única com um outro elemento. Um exemplo de função é *ser_pais_biológicos*, onde um conceito *homem* e um conceito *mulher* estão relacionados a um conceito *pessoa*.
- Um conjunto de axiomas, ou seja, regras que são sempre verdade. Um exemplo de axioma é afirmar que toda *pessoa* tem uma mãe.
- Um conjunto de instâncias que são um conhecimento prévio existente na ontologia.

Na definição apresentada por Gómez-Pérez (1999) inclui-se também como fazendo parte da ontologia as instâncias, o que não está de acordo com a definição apresentada por Maedche (2002), onde as instâncias fazem parte da base de conhecimento. Consideramos que algumas instâncias de conceitos fazem parte da ontologia, instâncias essas que são inerentes ao próprio conhecimento prévio do domínio. Porém, a grande maioria das instâncias está na base de conhecimentos.

4.4. Classificação das ontologias

Existem várias classificações diferentes de ontologias fornecidas por diferentes autores (Maedche, 2002; Gómez-Pérez, 1999). A que nos parece mais interessante, por utilizar uma característica chave das ontologias, é o sistema de classificação que usa a conceitualização como o principal critério para classificação. Esse sistema é introduzido em (Maedche, 2002) por N. Guarino no seu artigo “*Formal ontology and information systems*” sugerindo o desenvolvimento de diferentes tipos de ontologia de acordo com o nível de generalidade necessária. A seguir, são apresentados os 4 tipos propostos de classificação de uma ontologia (Maedche, 2002):

- Ontologias de alto-nível – Descrevem conceitos muito gerais como espaço, tempo, evento, etc. Esses conceitos tipicamente são independentes de um problema particular ou domínio. Sendo assim, é bem razoável ter-se uma ontologia de alto-nível compartilhada por grandes comunidades de usuários.
- Ontologias de domínio – Descrevem o vocabulário relacionado a um domínio genérico, através da especialização de conceitos introduzidos nas ontologias de alto-nível. São exemplos de ontologia de domínio ontologias de veículos, documentos, etc.
- Ontologias de tarefa – Descrevem um vocabulário relacionado a uma tarefa ou atividade genérica, através da especialização de conceitos introduzidos nas ontologias de alto-nível.
- Ontologias de aplicação – São as ontologias mais específicas por serem utilizadas dentro das aplicações. Esse tipo de ontologia especializa conceito tanto das ontologias de domínio, como também das de tarefas. Um exemplo é uma ontologia para uma aplicação que trabalhe com carros de luxo. Essa ontologia especializará conceito da ontologia de veículos (que é uma ontologia de domínio).

Através dessas descrições pode-se observar que as ontologias de alto-nível são as que possuem maior capacidade de reuso, por definir conceitos genéricos, enquanto as ontologias de aplicação são as que possuem menor capacidade de reuso, por definir conceitos relativos a uma aplicação específica.

A figura 13 mostra os diferentes tipos de ontologias e suas relações.

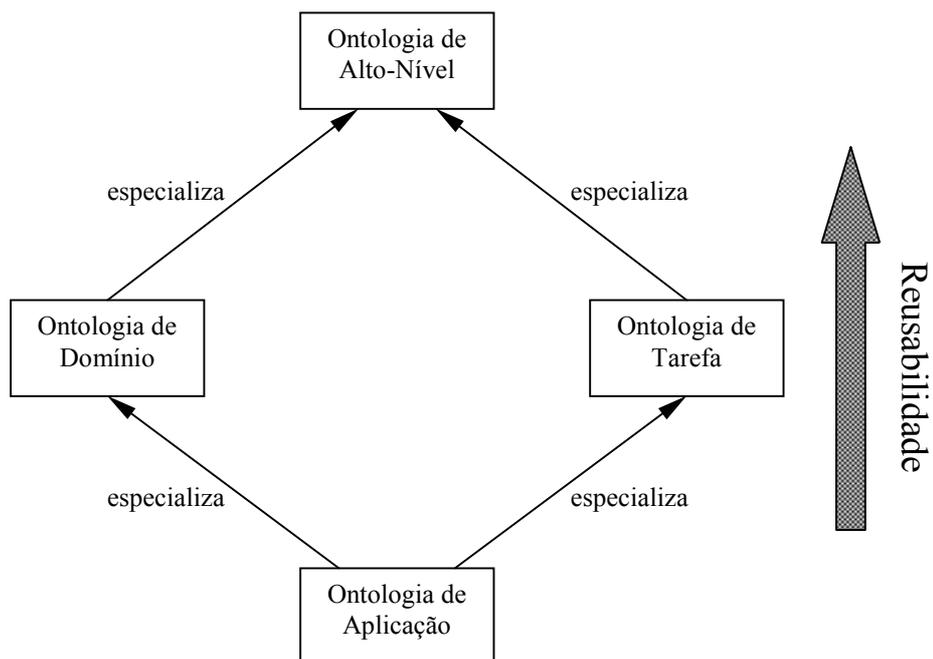


Figura 13 – Diferentes tipos de ontologias e seus relacionamentos.

4.5. **Principais áreas de aplicação das ontologias**

A seguir apresentaremos uma visão das várias áreas que empregam a tecnologia de ontologias.

4.5.1. **Processamento de linguagem natural**

Na tarefa de processar linguagem natural, o conhecimento do domínio é muito importante para uma compreensão coerente do texto. Esse conhecimento do domínio pode ser dado por meio de uma ontologia sobre o domínio de discurso do texto. A aplicação apresentada por Everett et al. (2002) é um exemplo da utilização de ontologias dentro da área de processamento de linguagem natural.

O uso de ontologias é de vital importância por dois motivos:

- Auxilia a elucidação de ambigüidades de compreensão existentes no texto. Com a utilização de uma ontologia sobre o domínio de discurso do texto se reduzem os problemas de ambigüidade.
- A ontologia funciona como um dicionário de conceitos dentro do domínio do texto.

4.5.2. **Gestão do conhecimento**

Empresas de consultoria como Arthur Andersen, Ernst & Young e Price Waterhouse têm como principal valor da empresa o conhecimento. Achar a informação certa o mais rápido possível dentro da empresa é de vital importância para a mesma.

Sistemas de gestão do conhecimento lidam com a aquisição, manutenção e acesso ao conhecimento dentro de uma organização. A tecnologia de ontologias dentro dessa área auxilia das seguintes formas:

- Ontologias fornecem a estrutura básica sobre a qual se constroem bases de conhecimentos.
- Uma dificuldade dos sistemas de gestão de conhecimento é o fato de que muito do conhecimento presente dentro das organizações se encontra em uma forma não estruturada. Usando ontologias, pode-se anotar informações semânticas em artefatos de informação não

estruturados, visando assim a obtenção de resultados mais precisos em pesquisas de informação.

4.5.3. Web semântica

Uma citação que caracteriza muito bem o estado atual da *Web* é a seguinte, de autoria de John Naisbitt em seu livro "*Megatrends 2000*" (Staab, 2002):

“Nós estamos nos afogando em informações e com fome de conhecimento”

Podemos ver claramente como essa citação é verdadeira quando pensamos nas dificuldades existentes em obter resultados precisos em buscas tipicamente realizadas na *Web*. Conforme vimos no capítulo 3, as ferramentas de busca que existem atualmente não conseguem descobrir o significado preciso sobre o que trata determinada página na *Web*, tendo que contar com certas heurísticas para tentar compreender o significado do conteúdo da página. Por exemplo, uma heurística muito utilizada é tentar classificar a página segundo a frequência de ocorrência das palavras na mesma.

A proposta da *Web* semântica feita por Tim Bernes-Lee (Berners-Lee et al., 2001) vem resolver esse tipo de problema adicionando semântica às páginas da *Web*. Para isso ele conta com três tecnologias principais: XML, RDF e ontologias. A tecnologia de ontologias entra para fornecer uma estrutura semântica para anotação das páginas da *Web*. Espera-se que com a estrutura fornecida pela *Web* semântica seja possível obter buscas mais precisas (uma vez que a semântica estará colocada em linguagem formal) e dar uma maior capacidade para os agentes de software que utilizem conteúdo da *Web*.

4.5.4. Comércio eletrônico

Tanto a área de *Business to Business* como também a área de *Business to Consumer* pode tirar proveito do uso de ontologias.

Na área de *Business to Business*, a automatização de transações requer uma descrição formal de produtos além de formatos de trocas sintáticas. Um entendimento comum dos termos e suas interpretações são capturados na forma de uma ontologia, permitindo assim interoperabilidade e meios para uma integração inteligente de informações (Maedche, 2002).

Já na área de *Business to Consumer*, o uso de ontologias vem para solucionar as dificuldades existentes na construção de *shopbots*. Uma vez que

todas as lojas on-line sigam uma mesma ontologia para a descrição de seus produtos, a tarefa de integração de catálogos das diversas lojas fica muito mais fácil. Ainda que não haja a adoção de uma ontologia única para a descrição de produtos a tarefa de mapeamento entre ontologias é muito mais simples em comparação com as dificuldades atuais. Uma discussão mais aprofundada a respeito das dificuldades atuais para a construção de *shopbots* pode ser vista na seção 3.3.1.

4.6.

Abordagens para o desenvolvimento de uma ontologia

Assim como na atividade de desenvolvimento de um software, o desenvolvimento de uma ontologia pode ser feito a partir de várias abordagens, cada um apresentando suas vantagens e suas desvantagens.

Conforme o caso deve-se escolher a abordagem mais adequada, segundo a classe da ontologia a ser desenvolvida.

A seguir são apresentadas as principais abordagens para o desenvolvimento de uma ontologia presente em (Holsapple & Joshi, 2002).

4.6.1.

Abordagem partindo de uma inspiração

Nessa abordagem, o desenvolvedor começa com uma premissa sobre por que a ontologia é necessária. Usando sua imaginação, criatividade e visão pessoal sobre o domínio de interesse da ontologia, o desenvolvedor constrói a ontologia que visa solucionar essa necessidade encontrada. Resumidamente, encontra-se uma ontologia para solucionar um problema específico.

A principal vantagem dessa abordagem está no fato de ser possível obter uma ontologia inovadora para a resolução de um problema. A principal desvantagem dessa abordagem é que a ontologia gerada é resultado de uma visão muito subjetiva, do próprio desenvolvedor sobre o domínio, gerando assim dificuldades para a sua adoção.

4.6.2.

Abordagem indutiva

Nessa abordagem, a ontologia é desenvolvida através da observação, exame e análise de casos específicos no domínio de interesse da ontologia. A ontologia resultante para um caso específico é aplicada para todos os outros casos do mesmo domínio.

A principal vantagem dessa abordagem é ser possível utilizar uma mesma ontologia para solucionar uma série de casos dentro de um domínio. Porém a principal dificuldade é a incapacidade dessa ontologia em cobrir domínios menos específicos, ou seja, mais amplos.

4.6.3. Abordagem dedutiva

Nessa abordagem, a ontologia é construída adotando-se alguns princípios gerais e aplicando-se adaptativamente estes princípios para construir uma ontologia específica. Isso envolve filtrar e destilar as noções gerais a fim de que elas possam ser personalizadas para um subconjunto do domínio (Holsapple & Joshi, 2002).

Em outras palavras, nessa abordagem, a partir de uma ontologia mais genérica de um domínio, chega-se a uma ontologia mais específica e restritiva dentro de um subconjunto do domínio. Por exemplo, partindo-se de uma ontologia genérica do domínio de automóveis, pode-se construir uma ontologia mais específica do domínio de automóveis utilitários que é um subconjunto domínio de automóveis.

A vantagem dessa abordagem está na capacidade de utilizar ontologias genéricas para geração de ontologias específicas. Porém a desvantagem é a necessidade de que já exista inicialmente uma ontologia genérica.

4.6.4. Abordagem sintética

Nessa abordagem, o desenvolvedor identifica um conjunto base de ontologias, em que nenhuma substitua a outra. As características dessas ontologias base, juntamente com outros conceitos pertencentes ao domínio que é descrito, são sintetizadas a fim de desenvolver uma ontologia unificada. Dessa forma, nessa abordagem o principal papel do desenvolvedor é fazer, de uma forma coerente, a composição de diversas ontologias em uma ontologia unificada.

Uma discussão interessante a respeito da unificação (composição) de ontologias pode ser vista em (Stephens & Huhns, 2001). Nesse é proposto uma ferramenta que auxilia o desenvolvedor no processo de composição de ontologias.

Por unir múltiplas ontologias, essa ontologia unificada tem grande propensão a ser aceita pelos usuários das ontologias usadas para a unificação. Dessa forma, essa ontologia unificada fornece uma forma coerente para a comunicação entre esses usuários das ontologias múltiplas.

A dificuldade dessa abordagem está na dificuldade muitas vezes existente para estabelecer uma ligação de forma harmônica entre um conjunto grande de ontologias.

4.6.5.

Abordagem colaborativa

Na abordagem colaborativa, o desenvolvimento é um esforço refletindo a experiência e os pontos de vistas das pessoas que intencionalmente cooperam para produzir a ontologia, ou seja, nessa abordagem o processo de desenvolvimento da ontologia não se encontra ao encargo de uma única pessoa e sim de um conjunto de pessoas.

A vantagem da abordagem colaborativa de desenvolvimento está em aumentar as chances da ontologia ser aceita como um padrão para um determinado domínio, pois possui a contribuição de um conjunto de pessoas para a sua formação. Outro ponto importante é o fato de a ontologia resultante desse processo ser muito mais elaborada do que nas abordagens supracitadas devido à mesma absorver o ponto de vista de mais de um indivíduo.

A principal dificuldade nessa abordagem está em coordenar o grupo de desenvolvimento, quando esse é muito grande.

4.7.

Metodologias para o desenvolvimento de uma ontologia

O processo de construção de ontologias ainda se encontra pouco desenvolvido. A grande maioria dos desenvolvedores usa seus próprios critérios no processo de desenvolver uma ontologia.

Devido a essa falta de uma metodologia estabelecida para o desenvolvimento, uma prática comum entre os desenvolvedores de ontologias é passar diretamente do passo de aquisição de conhecimento para o passo de implementação, o que gera os seguintes problemas:

- Os modelos conceituais da ontologia ficam implícitos no código da implementação.
- Dificuldades de reuso da ontologia, pois o *design* da ontologia e as decisões de projeto estão implícitos no código.
- Gera problemas de comunicação devido às dificuldades que o *expert* no domínio da ontologia tem para entender o código da implementação. Isso é um sério problema, pois ele tende a ser a principal fonte de informação sobre o domínio.
- Gera dificuldades no desenvolvimento de ontologias complexas, pois a passagem da aquisição de conhecimento para a implementação é muito abrupta.
- Dependendo da linguagem escolhida para a codificação pode-se limitar a capacidade de descrição conceitual do domínio da ontologia.

Dessa forma, se faz necessária a adoção de uma metodologia para assim reduzirmos as dificuldades acima citadas.

Para esse trabalho analisamos três metodologias presentes na literatura e por fim escolhemos a metodologia que pareceu mais apropriada para ser utilizada no desenvolvimento da nossa ontologia. As metodologias analisadas foram às propostas em (Uschold & King, 1995; Grüninger & Fox, 1995; Fernández et al., 1997), sendo que essa última foi a escolhida.

A seguir apresentamos uma descrição rápida das três metodologias e falamos um pouco mais profundamente da metodologia escolhida.

4.7.1.

Metodologia proposta por Mike Uschold e Martin King (Uschold & King, 1995)

Essa metodologia compreende os seguintes estágios para o desenvolvimento de ontologias:

- Identificação do propósito – Visa identificar o porquê da construção da ontologia e as suas intenções de uso.
- Construção da ontologia – Esse estágio é subdividido em três estágios:
 - Captura da ontologia – Visa identificar conceitos e relacionamentos do domínio de interesse para produzir uma definição precisa dos mesmos.
 - Codificação – Codificar a ontologia em uma linguagem formal.
 - Integração com ontologias existentes – Integrar a nova ontologia com as ontologias existentes.
- Avaliação – Avaliação da ontologia.
- Documentação – Documentação da ontologia.

A principal desvantagem dessa metodologia é que ela não descreve de uma forma precisa as técnicas para execução das diferentes atividades. O nível de detalhamento da metodologia é muito pequeno, só oferecendo princípios gerais muito vagos.

4.7.2.

Metodologia proposta por Michael Grüninger e Mark S. Fox (Grüninger & Fox, 1995)

Essa metodologia foi desenvolvida baseada na experiência dos autores no desenvolvimento de ontologias para empresas. Ela é formada pelos seguintes estágios:

- Definir os cenários motivadores – Identifica possíveis problemas que demandem uma nova ontologia. O cenário motivador também fornece intuitivamente um conjunto de soluções possíveis para o problema.
- Definir informalmente questões de competência – Dado o cenário motivador, um conjunto de perguntas irão surgir que necessitarão de

uma ontologia para que elas sejam respondidas. Essas perguntas são as questões de competência da ontologia. Elas não são expressas em linguagem formal.

- Especificação em lógica de primeira ordem da terminologia – Uma vez que foram definidas informalmente as questões de competência a fim de propor ou estender uma ontologia, a terminologia da ontologia deve então ser especificada usando lógica de primeira ordem ou equivalente.
- Especificar as questões de competência formalmente – Uma vez que foram definidas informalmente as questões de competência e a terminologia da ontologia, as questões de competência são definidas em linguagem formal.
- Especificação em lógica de primeira ordem dos axiomas – Os axiomas na ontologia especificam definições de termos da ontologia e limitações de sua interpretação. Esses axiomas são definidos em lógica de primeira ordem usando os predicados da ontologia.
- Verificação através dos teoremas de completude – Através desses teoremas são fornecidos meios de determinar a extensibilidade da ontologia, fazendo explicitamente o papel que cada axioma executa no teorema.

A figura 14 apresenta o procedimento para o desenvolvimento e avaliação de ontologia.

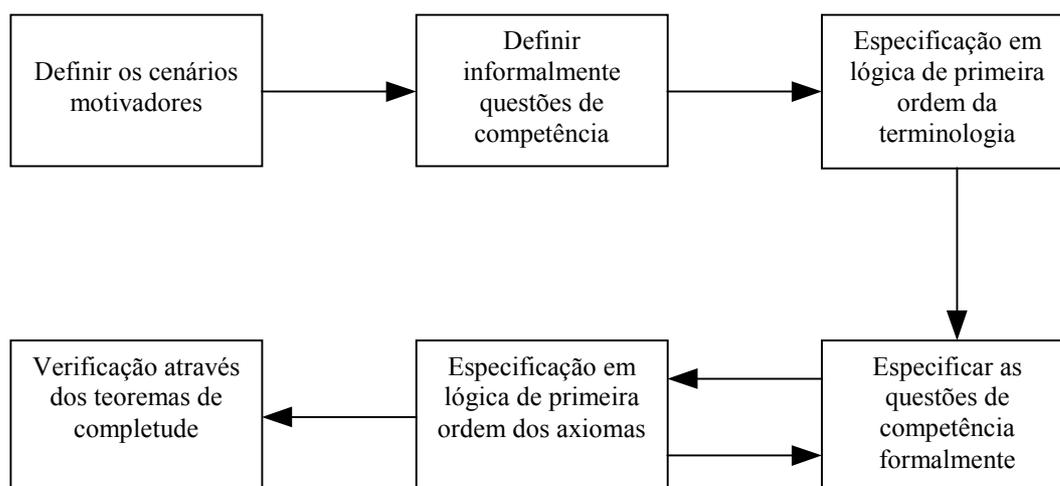


Figura 14 - Procedimento para o desenvolvimento e avaliação de ontologia.

Diferentemente da metodologia anterior, essa fornece mais do que princípios gerais. Olhando o exemplo de utilização da ontologia apresentado no

artigo de Grüninger & Fox (1995), podemos ver que a partir da segunda etapa da metodologia (Definir informalmente questões de competência) a metodologia exige uma codificação em linguagem formal. Em outras palavras, após a aquisição do conhecimento sobre o domínio do problema que vai até a segunda etapa, a metodologia parte para a geração de artefatos em lógica de primeira ordem, ou equivalente.

Apesar de esse formalismo ser adequado para avaliar se a ontologia atende os requisitos e evitar ambigüidades na especificação, ele dificulta muito a comunicação entre o desenvolvedor da ontologia e o *expert* do domínio que provavelmente é um componente muito importante no desenvolvimento de ontologias. Sendo assim, a desvantagem dessa metodologia está na falta de modelos intermediários que facilitem a comunicação entre desenvolvedor da ontologia e o *expert* do domínio.

4.7.3.

Metodologia proposta por Mariano Fernández, Asunción Gómez-Pérez e Natalia Juristo (METHONTOLOGY) (Fernández et al., 1997)

Dos três artigos analisados que propõem metodologias para o desenvolvimento de ontologias, este parece ser o mais maduro de todos, pois além de descrever mais a fundo a metodologia no que tange os passos a serem seguidos e dos artefatos a serem criados para a geração de um modelo conceitual, fornece também um processo de desenvolvimento de ontologias e também propõe um ciclo de vida baseado em evolução de protótipos.

O processo de desenvolvimento de ontologias proposto identifica que as seguintes atividades devem ser cumpridas quando da construção de uma ontologia:

- Planejamento – Identifica as tarefas principais da ontologia, e planeja a utilização dos recursos.
- Especificação – Define por que a ontologia está sendo construída, e quem serão seus usuários.
- Aquisição de Conhecimento – Adquire conhecimento sobre o domínio da ontologia. Pode ser feito: através de entrevistas com *expert* no domínio da ontologia a ser criada, análise de livros sobre o domínio, etc.

- Conceitualização – Criação de um modelo conceitual que descreve o problema e sua solução.
- Formalização – Transforma o modelo conceitual em um modelo formal ou semiformal.
- Integração – Procura integrar o máximo possível as ontologias existentes à nova ontologia.
- Implementação – Implementa a ontologia em uma linguagem formal de modo que ela seja computável.
- Avaliação – Avalia a ontologia.
- Documentação – Faz a documentação da ontologia visando facilitar futuro reuso e manutenção.
- Manutenção – Executar a manutenção da ontologia quando necessária.

A figura 15 apresenta os estágios do ciclo de vida e as atividades propostas. Conforme essa figura, as atividades de aquisição de conhecimento, avaliação e documentação, são executadas em todos os estágios do ciclo de vida. O planejamento é a primeira atividade a ser executada. A maior parte da atividade de aquisição é feita simultaneamente com o estágio de especificação da ontologia e vai decaindo conforme o ciclo de vida avança. A maior parte da atividade de avaliação da ontologia é feita durante os estágios iniciais do ciclo, de forma a diminuir a propagação de erro. A atividade de documentação deve ser realizada em todos os estágios.

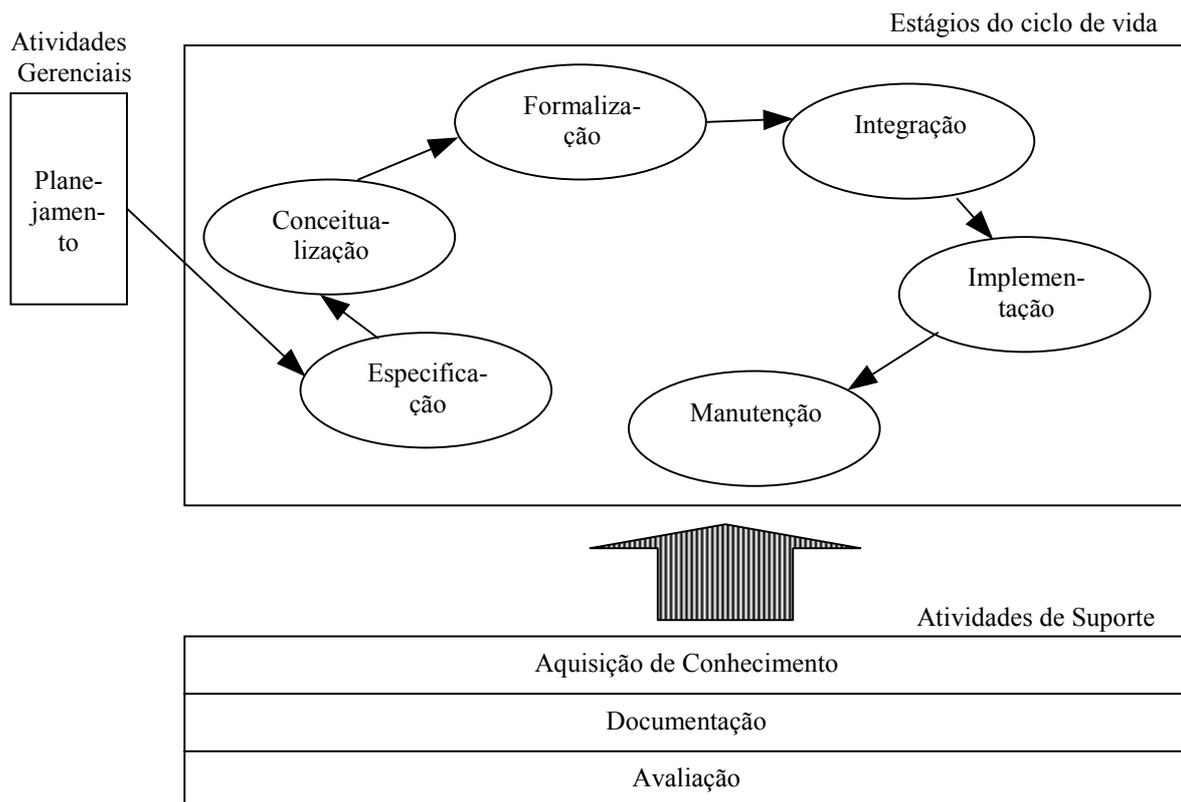


Figura 15 – Visão geral do processo de desenvolvimento de ontologias (Fernández et al., 1997).

A parte que mais nos interessa aqui na metodologia é onde são descritos os artefatos que devem ser gerados no seu desenvolvimento, principalmente aqueles gerados na parte de conceitualização da ontologia. Esses artefatos não são codificados em lógica formal e sim em uma representação intermediária que descreve a ontologia. Assim sendo, esses artefatos são de fácil entendimento tanto para o desenvolvedor da ontologia quanto para o *expert* no domínio. Outra grande vantagem da geração desses artefatos durante a conceitualização é que eles fornecem uma boa documentação sobre a ontologia. Uma discussão mais aprofundada sobre os artefatos pode ser visto em (López et al., 1999), onde o autor constrói uma ontologia sobre o domínio da química utilizando essa metodologia.

O artefato proposto na metodologia que é gerado durante o estágio de especificação é um documento que cobre o objetivo principal, o propósito, a granularidade e o escopo da ontologia. O objetivo é de identificar o conjunto de termos a serem representados, suas características, e sua granularidade. Essa especificação deve ser o mais completa e concisa possível.

Os principais artefatos propostos na metodologia que são gerados durante o estágio de especificação e conceitualização são (López et al., 1999):

- Glossário de termos – Inclui todos os termos do domínio (conceitos, instâncias, atributos, verbos, etc.) e suas descrições. A descrição deve ser clara e concisa.
- Árvore de classificação de conceitos – Define relações tais como: subclasse e exclusividade mútua entre classes. Dessa forma são definidas várias taxinomias do domínio, sendo que cada uma gera uma ontologia. Deve-se verificar a consistência (não devem existir ciclos na árvore) e a concisão (não deve existir repetição de conceitos).
- Diagrama de relações binárias – Esse diagrama estabelece os relacionamentos entre conceitos da mesma ontologia e de ontologias diferentes.
- Dicionário de conceitos – Esse documento contém todos os conceitos do domínio, instâncias desses conceitos, atributos de classes e atributos de instâncias dos conceitos, e opcionalmente, sinônimos e antônimos dos conceitos. Para cada árvore de classificação de conceitos deve existir um dicionário de conceitos.
- Tabela de relações binárias – Especifica o nome da relação, o nome dos conceitos origem e destino da relação, a relação inversa e a cardinalidade da relação. Para cada árvore de classificação de conceitos deve existir uma tabela de relações binárias.
- Tabela de atributos de instância – Descreve cada atributo de instância do dicionário de conceitos. Atributos de instância são os atributos definidos nos conceitos, mas que são valorados nas instâncias. Para cada atributo de instância deve ser especificado: nome, tipo de valor, unidade de medida para valores numéricos, precisão dos valores numéricos, faixa de valores aceitos, valor padrão, cardinalidade máxima e mínima, atributos ou constantes utilizadas para inferir o valor do atributo que está sendo definido, atributos que podem ser inferidos desse atributo, fórmulas ou regras para inferir o atributo que está sendo definido e por fim as referências usadas para preencher o atributo.

- Tabela de atributos de classe – Descreve os atributos de classe no dicionário de conceitos. Esses atributos são aqueles em que o valor é dado para o conceito, ou seja, o valor será o mesmo para todas as instâncias do conceito. Para cada atributo de classe deve ser especificado: nome do atributo, tipo de valor, unidade de medida para valores numéricos, precisão dos valores numéricos, cardinalidade máxima e mínima, atributos que podem ser inferidos desse atributo, e referências.
- Tabela de axiomas – Define conceitos por meio de expressões lógicas que são sempre verdadeiras. Para cada axioma deve ser definido: nome, descrição em linguagem natural, os conceitos aos quais o axioma se refere, os atributos usados no axioma, a expressão lógica que define o axioma formalmente e referências.
- Tabela de constantes – Especifica para cada constante: nome, descrição em linguagem natural, o tipo lógico do valor (o que o valor é), seu valor constante, sua unidade de medida, os atributos que podem ser inferidos usando a constante e referências.
- Tabela de fórmulas – Descreve cada fórmula das tabelas de atributos de instância. Essa tabela é usada para inferir os valores numéricos dos atributos de instância. Para cada fórmula deve ser especificado: nome, atributos inferidos com a fórmula, expressão matemática, descrição em linguagem natural, atributos e constantes usados no cálculo da fórmula, precisão do cálculo, situações em que a fórmula faz sentido e referências.
- Árvore de classificação de atributos – Mostra graficamente os atributos e as constantes relacionados a uma seqüência de cálculo de um atributo raiz. É usado para validar que todos os atributos usados na fórmula fazem sentido e nenhum atributo foi omitido.
- Tabela de instâncias – Descreve as instâncias do domínio. É descrito através do nome, seus atributos de instância e valores para os mesmos.

A figura 16 apresenta os principais artefatos gerados durante a execução da metodologia.

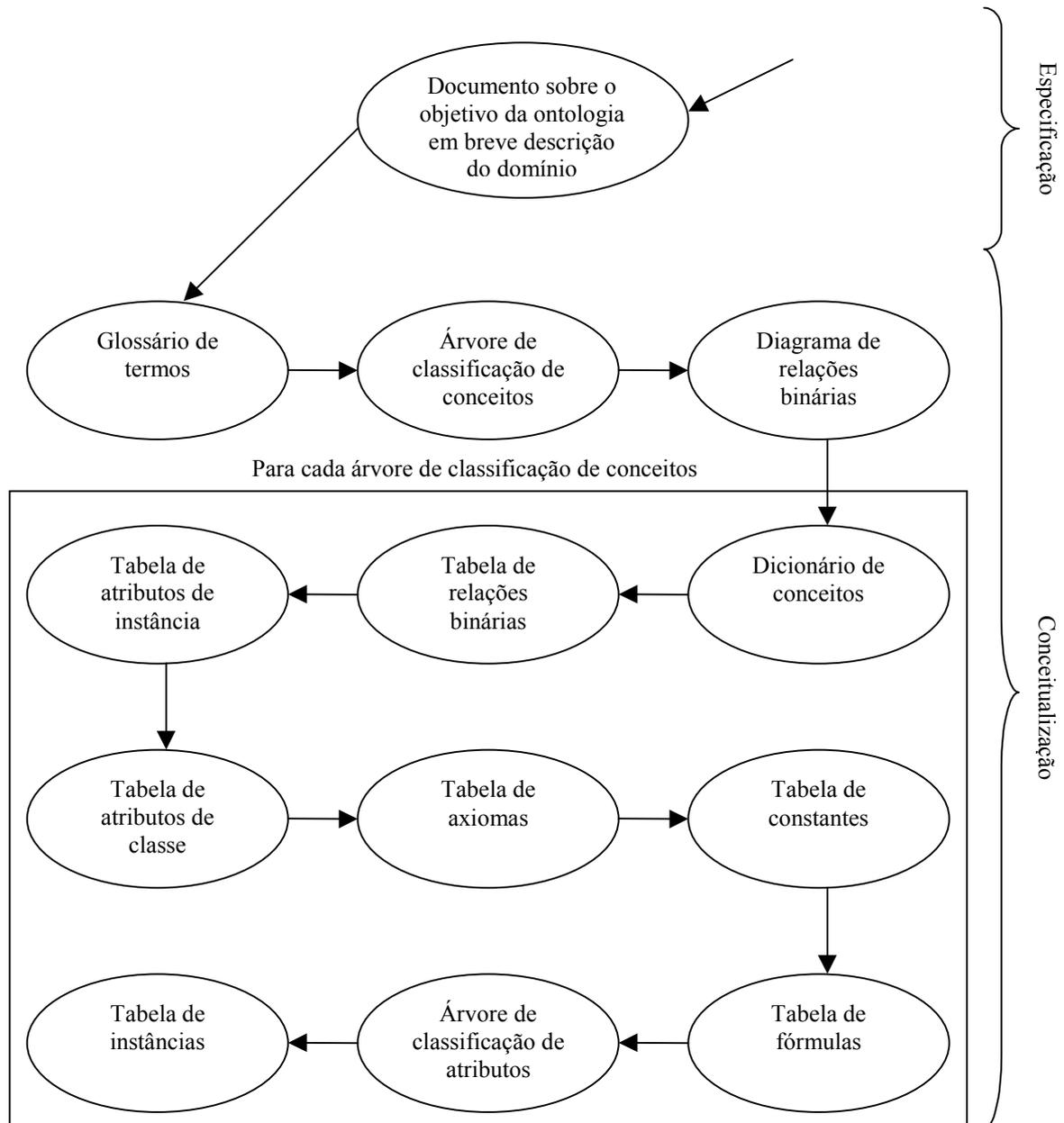


Figura 16 – Artefatos gerados na metodologia proposta em (Fernández et al., 1997; López et al., 1999).

4.7.4. Conclusão a respeito das metodologias

A metodologia METHONTOLOGY foi a escolhida para a modelagem conceitual da ontologia utilizada nesse trabalho. Isso se deve principalmente à sua clareza na descrição dos artefatos a serem gerados na descrição conceitual da ontologia.