

**Ontologias: conceitos, usos, tipos, metodologias,
ferramentas e linguagens**

E. A. M. Moraes A. P. L. Ambrósio

Technical Report - INF_001/07 - Relatório Técnico
December - 2007 - Dezembro

The contents of this document are the sole responsibility of the authors.
O conteúdo do presente documento é de única responsabilidade dos autores.

Ontologias: conceitos, usos, tipos, metodologias, ferramentas e linguagens

Edison Andrade Martins Morais *

edison@inf.ufg.br

Ana Paula L. Ambrósio †

apaula@inf.ufg.br

Abstract. *The objective of this technical report is describe the ontologie concepts, its main uses, types, development methodologies, specification tools and representation languages.*

Keywords: Ontologie.

Resumo. *O objetivo deste relatório técnico é descrever os conceitos de ontologia, seus principais usos, tipos, metodologias de desenvolvimento, ferramentas de especificação e linguagens de representação.*

Palavras-Chave: Ontologia.

1 Introdução

Atualmente, o volume de dados disponíveis, principalmente na Internet, é muito grande, e vem crescendo rapidamente [29]. Devido a estas altas taxas de crescimento, torna-se necessária a utilização de técnicas para melhoria no tratamento e organização destes dados, atuando principalmente na sua seleção, processamento, recuperação e disseminação [2].

Segundo Almeida et al. [2], técnicas de tratamento e organização de dados podem ser classificadas de diversas formas, por exemplo, a partir de seus termos, em *glossários* ou *dicionários*, por classificação ou categorias, através de *taxonomias*, ou a partir de conceitos e seus relacionamentos, utilizando *ontologias*, *tesauros* ou *redes semânticas*.

Ontologia é uma técnica de organização de informações que vem recebendo especial atenção nos últimos anos, principalmente no que diz respeito à representação formal de conhecimento [22] [24]. Geralmente criadas por especialistas, tendo sua estrutura baseada na descrição de conceitos e dos relacionamentos semânticos entre eles, as ontologias geram uma especificação formal e explícita de uma conceitualização compartilhada [4].

Uma razão para a popularidade das ontologias existe devido à grande promessa de compartilhamento e entendimento comum de algum domínio de conhecimento que possa ser comunicado entre pessoas e computadores. Neste sentido, ontologias têm sido desenvolvidas para facilitar o compartilhamento e reutilização de informações [21].

*Mestrando em Ciência da Computação – INF/UFG.

†Orientadora – INF/UFG.

A comunidade da Internet acredita que, em um futuro próximo, todo negócio da rede deverá fornecer a semântica de suas páginas através de uma ontologia [15], pois as ontologias oferecem uma “Língua Franca” que permite que máquinas processem e integrem recursos de Internet de maneira inteligente, possibilitando buscas rápidas e acuradas e facilitando a comunicação entre dispositivos heterogêneos na rede.

O objetivo deste relatório técnico é descrever os conceitos de ontologia, seus principais usos, tipos, metodologias de desenvolvimento, ferramentas de especificação e linguagens de representação.

Este relatório está organizado da seguinte forma: a Seção 2 descreve o conceito de ontologia em seu sentido filosófico e relacionado à ciência da computação, seus principais usos, benefícios e problemas, os tipos de ontologias existentes, e uma forma de construí-las; A Seção 3 descreve os principais usos, benefícios e problemas relacionados às ontologias; A Seção 4 descreve os tipos de ontologias existentes; A Seção 5 descreve as principais metodologias, ferramentas e linguagens utilizadas na Construção de Ontologias; A conclusão está descrita na Seção 6.

2 Ontologia

2.1 Sentido Filosófico

Em seu sentido filosófico, ontologia possui diversas definições. De acordo com Almeida et al. [2], é o ramo da metafísica¹ que estuda os tipos de coisas que existem no mundo. A palavra é derivada do grego *ontos*, ser, e *logos*, palavra. Entretanto, seu termo de origem é a palavra aristotélica *categoria*, termo este utilizado no sentido de classificação. Neste sentido, Aristóteles² apresenta categorias que servem de base para classificar entidades, e introduz o termo *differentia* (Figura 1) para propriedades diferentes na classificação de espécies de mesmo gênero (categoria).

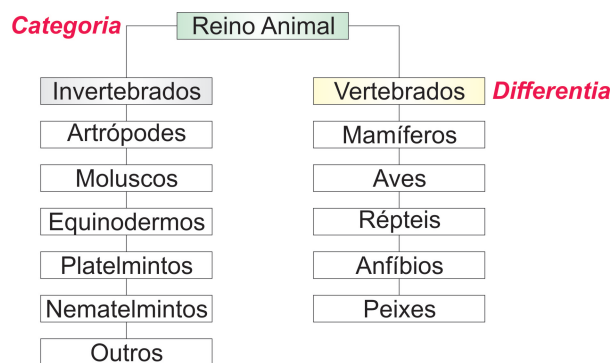


Figura 1: Ontologia no Sentido Filosófico

Segundo Blackburn et al. [5], é também a parte da filosofia que trata da natureza do ser, ou seja, da realidade, da existência dos entes e das questões metafísicas em geral. Já Smith [45],

¹É um ramo da filosofia que estuda o mundo como ele é, ou seja, é o estudo do ser ou da realidade. É a ciência do ser como ser, ou dos princípios e das causas do ser e de seus atributos essenciais. Ela abrange ainda o ser imóvel e incorpóreo, princípio dos movimentos e das formas do mundo, bem como o mundo mutável e material, mas em seus aspectos universais e necessários [5].

²Filósofo grego nascido em Estagira, colônia grega da Trácia, no litoral setentrional do mar Egeu, em 384 a.C, foi um dos maiores pensadores de todos os tempos e considerado o criador do pensamento lógico [5].

refere-se ao termo como a ciência do “o que é?” ou o “estudo das coisas que existem”, da descrição de tipos e estruturas de objetos, suas propriedades, eventos, processos e relacionamentos em cada área do mundo real.

Entretanto, apesar dos filósofos utilizarem ontologias durante anos para descrever domínios naturais do mundo, ainda não há um consenso, principalmente na área de Ciência da Computação, sobre a real semântica do termo ontologia [25].

2.2 Ontologias e Ciência da Computação

Existem na literatura várias definições para o termo ontologia em Ciência da Computação. Gruber [21] afirma que ontologia é uma especificação de uma conceitualização, ou seja, é uma descrição de conceitos e relacionamentos que existem entre estes conceitos. Esta definição genérica, bastante diferente do sentido filosófico de ontologia, considera as ontologias apenas como um conjunto de conceitos e definições.

Uma definição mais específica pode ser encontrada em [7]. Neste caso, uma ontologia é definida como uma especificação *formal e explícita* de uma *conceitualização compartilhada*, onde especificação formal quer dizer algo que é legível para os computadores, explícita são os conceitos, propriedades, relações, funções, restrições e axiomas explicitamente definidos, conceitualização representa um modelo abstrato de algum fenômeno do mundo real e compartilhada significa conhecimento consensual.

Entretanto, apesar de apresentarem diferentes definições, tanto [21], quanto [7], afirmam que o principal propósito da construção de ontologias é permitir compartilhamento e reutilização de conhecimento.

Logo, para permitir compartilhamento e reutilização de conhecimento, em Ciência da Computação, é importante que os conceitos, presentes nas ontologias possuam de fato uma especificação formal. Neste sentido, [24] sugere uma maneira de criar definições formais para conceitos.

Seja a estrutura $\langle D, W \rangle$ tal que D representa o domínio em questão e W representa todos os conceitos existentes em D . Esta estrutura é chamada de Espaço de Domínios. Uma conceitualização C é uma estrutura $\langle D, W, R \rangle$ sendo R o conjunto de relações escolhidas como pertinentes para representar este domínio. Desta forma, uma conceitualização define uma estrutura pretendida do mundo, representada por S .

Do ponto de vista computacional, para que uma conceitualização possa ser efetivamente usada, ela precisa ser especificada em uma determinada linguagem L . A estrutura de S é, então, mapeada para constantes e predicados da linguagem L , seguindo uma função de interpretação.

Portanto, no sentido definido por [24], as ontologias passam a ser mecanismos de especificação parcial, tornando-se representativas apenas em relação a um determinado domínio e não com a completude do conhecimento. Ao conjunto destes elementos de um domínio que podem ser representados em uma ontologia é dado o nome de *Universo de Discurso* [25].

Na prática, uma ontologia define uma “linguagem”, conjunto de termos, que será utilizada para formular consultas [2]. A ontologia define as regras de combinação entre os termos e seus relacionamentos, estes relacionamentos são criados por especialistas, e os usuários formulam consultas usando os conceitos especificados.

3 Principais Usos, Benefícios e Problemas relacionados às Ontologias

Ontologias podem ser utilizadas em várias áreas da Ciência da Computação, dentre elas:

- *Recuperação de informações na Internet*

Projetos relacionados à recuperação de informação na Internet, dentre eles, *OntoSeek* [6], que recupera informações de catálogos de produtos on-line utilizando um sistema de agentes inteligentes, o *Ontoweb* (<http://www.ontoweb.com.br>), que é um sistema de análise de informações na Internet, que possibilita uma pesquisa contextualizada nas fontes acessadas, o SEAL (*Semantic Portal*) [35], que permite o desenvolvimento de portais semânticos baseados em ontologia, ou o *OntoSearch* [42], que permite a reutilização de ontologias na Web-Semântica [4] provendo estruturas de busca em um grande banco de dados de ontologias e outros documentos semânticos na Internet.

- *Processamento de Linguagem Natural*

Projetos de processamento de linguagem natural [19] [1] [2], como o *Oncoterm* [38], por exemplo, que facilita a tradução de textos médicos sobre oncologia mediante uma ontologia baseada em textos especializados e dicionários médicos.

- *Gestão do Conhecimento*

Projetos de representação (gestão) do conhecimento [12] [13] [2], como o CoMMA [17] (*Corporate Memory Management through Agents*), por exemplo, que propõe o armazenamento da memória corporativa da empresa através do uso de ontologias.

- *Web-Semântica*

A Web-Semântica é uma visão para o futuro da Internet na qual a informação é dada com significado explícito, tornando mais fácil para máquinas processarem e integrarem informações disponíveis na rede automaticamente [4].

- *Educação*

Existem alguns projetos na área de educação em ambiente de Internet, por exemplo, *RichODL* [43], que é um ambiente de aprendizado desenvolvido para treinar estudantes na modelagem e simulação de ambientes dinâmicos. Neste exemplo, as ontologias são utilizadas para descrever o domínio físico do sistemas modelados.

Já o *SchoolOnto* (*Scholarly Ontologies Project*) [44] é um biblioteca digital baseada em ontologias. Ela possibilita a interpretação de domínios, e auxilia na modelagem de pesquisas dinâmicas que necessitam de ferramentas para tratar inconsistências.

Além de serem utilizadas em diversas áreas, o uso de ontologias traz diversas vantagens. De acordo com Guizzardi [25], os principais benefícios relacionados à utilização de ontologias são:

- *Comunicação*

As ontologias possibilitam a comunicação entre pessoas acerca de determinado conhecimento, pois permitem raciocínio e entendimento sobre um domínio. Essa relação auxilia na obtenção de consenso, principalmente sobre termos técnicos, entre comunidades profissionais, de pesquisa, etc.

- *Formalização*

A formalização está relacionada à especificação da ontologia, que permite eliminar contradições e inconsistências na representação de conhecimento, além de não ser ambígua. Além disso, essa especificação pode ser testada, validada e verificada.

- *Representação de Conhecimento e Reutilização*

As ontologias formam um vocabulário de consenso que permite representar conhecimento de um domínio em seu nível mais alto de abstração, possuindo, desta forma, potencial de reutilização.

Apesar de vários usos e benefícios, as ontologias ainda apresentam alguns problemas [25], dentre eles:

- *Escolha das Ontologias*

A escolha de uma ontologia pode se tornar um processo político, uma vez que uma ontologia pode não ser totalmente adequada a todos os indivíduos ou grupos relacionados a algum domínio específico.

- *Criação e Evolução das Ontologias*

Ontologias precisam ser criadas e evoluir. Poucos trabalhos, entretanto, têm focado este aspecto, principalmente na língua portuguesa. Alguns trabalhos sobre ontologias de domínio podem ser obtidos em [11].

- *Bibliotecas de Ontologias*

A noção de biblioteca de ontologias está relacionada à independência entre elas. Desta forma, a interface entre estas ontologias se constitui em um problema, uma vez que cada uma delas pode ter sido desenvolvida em um contexto diferente.

- *Metodologia de Desenvolvimento*

Este é considerado o principal problema relacionado às ontologias, principalmente pela falta de trabalhos descrevendo metodologias para seu desenvolvimento. Existe uma ausência de atividades padronizadas, ciclos de vida e métodos sistemáticos de desenvolvimento, assim como um conjunto de critérios de qualidade, técnicas e ferramentas. Isso faz com que o processo de criação de ontologias seja uma atividade quase “artística”.

Mais detalhes sobre como construir ontologias estão descritos na Seção 5.

4 Tipos de Ontologias

Ontologias podem ser classificadas quanto ao seu grau de formalismo, aplicação, conteúdo ou função (estrutura) [2].

Quanto ao grau de formalismo, as ontologias podem ser categorizadas em *altamente informais*, quando expressas em linguagem natural; *semi-informais*, quando expressas em linguagem natural de forma restrita e estruturada; *semi-formais*, expressas em linguagem artificial definida formalmente; e *rigorosamente formais*, onde os termos são definidos com semântica formal, teoremas e provas.

Em relação à aplicação, as ontologias podem ser de *autoria neutra*, quando, por exemplo, um aplicativo é descrito em uma única língua e depois convertido para o uso em diversos

sistemas, reutilizando-se as informações; como *especificação*, baseada em uma ontologia de domínio, que é utilizada para documentação e manutenção no desenvolvimento de softwares; e *de acesso comum a informação*, quando, por exemplo, um vocabulário é inacessível e a ontologia torna a informação possível de ser entendida, proporcionando conhecimento compartilhado dos termos.

Já em sua classificação quanto ao conteúdo, as ontologias podem ser *terminológicas*, quando representam termos que serão utilizados para modelar o conhecimento de um domínio específico; de *informação*, que especificam, por exemplo, a estrutura de registros de um banco de dados; de *modelagem de conhecimento*, que especificam as conceitualizações do conhecimento; de *aplicação*, que contém as definições necessárias para modelar conhecimento em uma aplicação; de *domínio*, que expressam conceitualizações que são específicas em um domínio; *genéricas*, que definem conceitos genéricos e comuns a várias áreas do conhecimento; e de *representação*, que explicam as conceitualizações que estão por trás dos formalismos de representação do conhecimento.

Quanto à sua função, as ontologias podem ser classificadas em cinco categorias [23] [25]:

- *Ontologias Genéricas*

São consideradas ontologias “gerais”. Descrevem conceitos mais amplos, como elementos da natureza, espaço, tempo, coisas, estados, eventos, processos ou ações, independente de um problema específico ou domínio particular. De acordo com Guizzardi [25], pesquisas enfocando ontologias genéricas procuram construir teorias básicas do mundo, de caráter bastante abstrato, aplicáveis a qualquer domínio (conhecimento de senso comum).

Exemplos de trabalhos utilizando este tipo de ontologia podem ser obtidos em [34] e [37]. Eles estão relacionados principalmente ao uso de ontologias em seu sentido filosófico de categorização e lingüística.

- *Ontologias de Domínio*

Descrevem conceitos e vocabulários relacionados a domínios particulares, tais como medicina ou computação, por exemplo. Este é o tipo de ontologia mais comum, geralmente construída para representar um “micro-mundo” [25]. Em [11] estão listados alguns trabalhos que utilizam este tipo de ontologia.

- *Ontologias de Tarefas*

Descrevem tarefas ou atividades genéricas, que podem contribuir na resolução de problemas, independente do domínio que ocorrem, por exemplo, processos de vendas ou diagnóstico. Sua principal motivação é facilitar a integração dos conhecimentos de tarefa e domínio em uma abordagem mais uniforme e consistente, tendo por base o uso de ontologias [25]. Exemplos de trabalhos nesta categoria podem ser obtidos em [40].

- *Ontologias de Aplicação*

Descrevem conceitos que dependem tanto de um domínio particular quanto de uma tarefa específica. Devem ser especializações dos termos das ontologias de domínio e de tarefa correspondentes (Figura 2). Estes conceitos normalmente correspondem a regras aplicadas a entidades de domínio enquanto executam determinada tarefa [25].

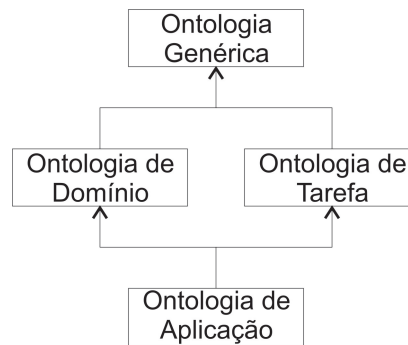


Figura 2: Ontologias de Aplicação [25]

- *Ontologias de Representação*

Explicam as conceituações que fundamentam os formalismos de representação de conhecimento, procurando tornar claros os compromissos ontológicos embutidos nestes formalismos [25]. Um exemplo desta categoria é a ontologia de *frames*, utilizada em Ontolândia [20].

Nas próximas seções estão descritas as principais metodologias, ferramentas e linguagens utilizadas na construção de ontologias de domínio.

5 Construção de Ontologias

A construção de ontologias de domínio envolve, primeiramente, a definição de seu domínio e escopo. Uma vez definidos estes elementos, devem ser escolhidos uma metodologia (Seção 5.1), uma ferramenta (Seção 5.2) e uma linguagem (Seção 5.3) para sua especificação.

Na prática, as metodologias, ferramentas e as linguagens são utilizadas na definição da estrutura da ontologia. Nem todas as ontologias têm a mesma estrutura [2], mas a maioria delas possui alguns elementos básicos, como:

- *Classes*

Normalmente organizadas em taxonomias, as classes representam algum tipo de interação da ontologia com um determinado domínio (Figura 3).

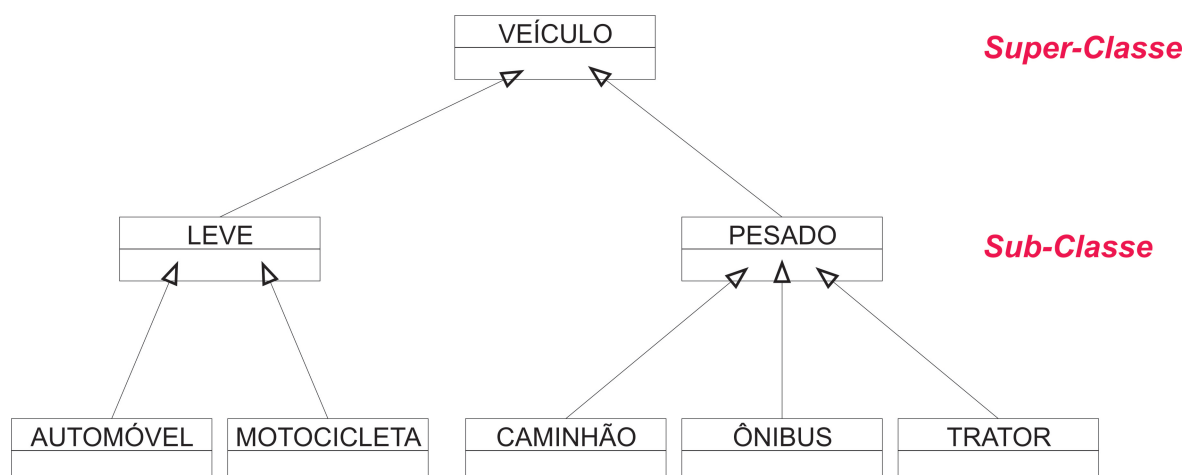


Figura 3: Exemplo de taxonomia de classes.

- *Relações*
Representam o tipo de interação entre os elementos do domínio (classes).
- *Axiomas*
São utilizados para modelar sentenças consideradas sempre verdadeiras [2].
- *Instâncias*
São utilizadas para representar elementos específicos, isto é, os próprios dados da ontologia.
- *Funções*
Eventos podem ocorrer no contexto da ontologia.

5.1 Metodologias para Construção de Ontologias

Metodologias de desenvolvimento de ontologias existem no intuito de sistematizar sua construção e manipulação. Entretanto, nenhuma das metodologias descritas a seguir é totalmente madura, principalmente se comparadas com metodologias de engenharia de software [2].

- *Enterprise* [30]

É baseada em quatro fases (Figura 4): identificação do propósito, identificação do escopo, formalização e documentação formal.

A *identificação do propósito* determina o nível de formalidade que a ontologia deve ser descrita; Na *identificação do escopo* uma especificação é produzida de acordo com o domínio que a ontologia precisa representar; A *formalização* é a criação do código, definições formais e axiomas relacionados à ontologia; A *documentação formal* é a fase onde a ontologia será documentada e as fases de identificação do escopo e formalização podem ser revistas.

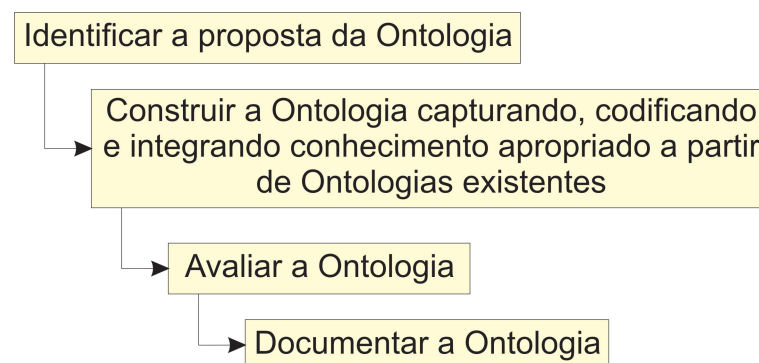


Figura 4: Metodologia *Enterprise*.

- *Methontology* [30]

É baseada na construção da ontologia a partir do conhecimento de um domínio. Suas atividades principais são (Figura 5): especificação de requisitos, conceitualização do domínio do conhecimento, formalização do modelo conceitual em uma linguagem formal, implementação de um modelo formal e manutenção de ontologias implementadas.

Esta metodologia possui ainda algumas atividades de suporte desempenhadas durante o processo de construção da ontologia: aquisição do conhecimento, integração, avaliação, documentação e gerenciamento de configuração.

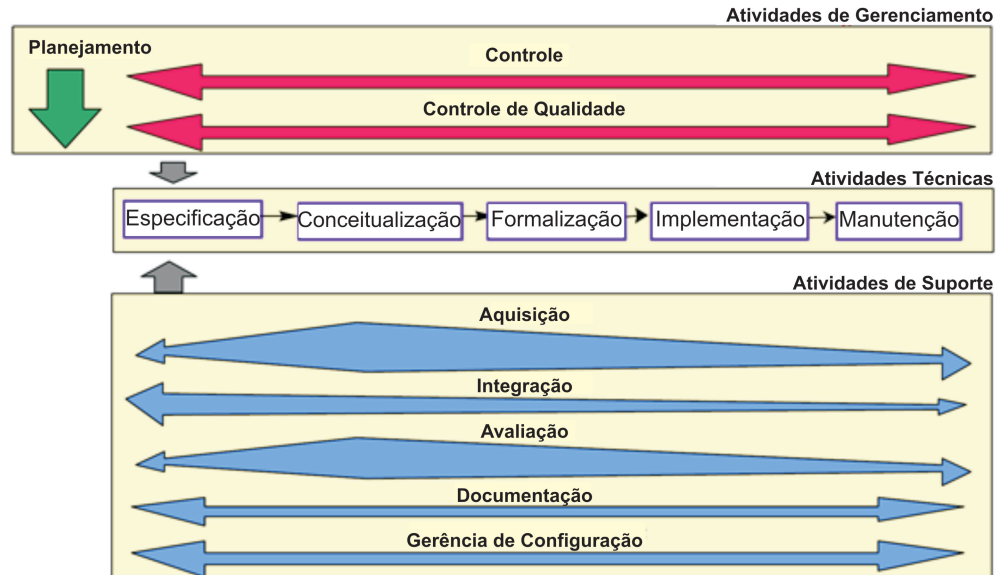


Figura 5: *Methontology*.

- *On-To-Knowledge* [46]

Esta metodologia é baseada em quatro fases (Figura 6): *kick-off*, refinamento, avaliação e manutenção.

No *kick-off* os requisitos para construção da ontologia são capturados e especificados, questões de competência são identificadas, ontologias potencialmente reutilizáveis são estudadas e uma primeira versão da ontologia é construída.

No *refinamento*, uma ontologia mais madura é construída a partir da primeira versão. Na *avaliação*, os requisitos e as questões de competência são checados e a ontologia é colocada em ambiente de produção. A *manutenção* envolve atividades de adaptação da ontologia às mudanças nos requisitos e correção de erros.

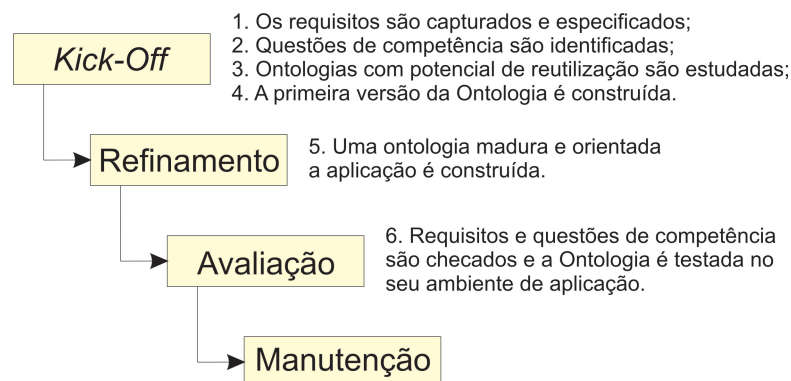


Figura 6: *On-to-Knowledge*.

Existem outras metodologias além das descritas nesta Seção. Mais informações sobre estas metodologias podem ser obtidas em [2] e [30].

Pelo fato das metodologias descritas nesta Seção ainda não demonstrarem um processo suficientemente estruturado a ponto de suportar a construção de ontologias como uma verdadeira disciplina de engenharia, Guizzardi [25] sugere uma abordagem sistemática para sua construção. Essa metodologia, inicialmente proposta por Falbo [14], é baseada em seis fases: identificação de propósito e especificação de requisitos, captura da ontologia, formalização da ontologia, integração com ontologias existentes, avaliação e documentação.

Ainda segundo Guizzardi [25], esta abordagem une as principais características das metodologias descritas nesta Seção. Seu ciclo de vida é iterativo, e alguns processos, como avaliação da qualidade e documentação, ocorrem durante todo o ciclo de vida (atividades guarda-chuva).

A figura 7 mostra o diagrama de atividades que ilustra esta metodologia.

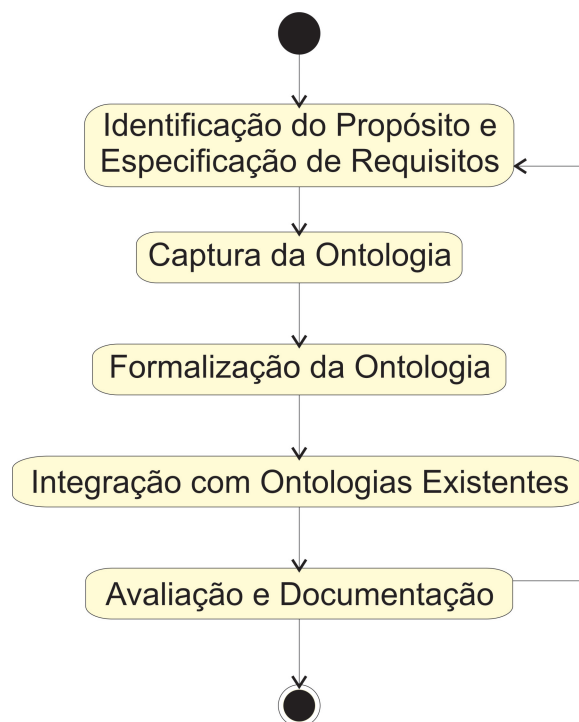


Figura 7: Processo Iterativo de Construção de Ontologias.

- *Identificação de propósito e especificação de requisitos*

Esta é a primeira atividade realizada na construção de ontologias. Seu principal objetivo é identificar a *competência da ontologia*, ou seja, seus usos e propósitos, através da delimitação do que é relevante para a ontologia e o que não é. Além disso, nesta fase, podem ser identificados potenciais usuários da ontologia, e o contexto que motivou sua construção.

Após definir a competência, finalmente os requisitos da ontologia devem ser especificados. Essa especificação envolve a descrição do propósito e dos usos da ontologia, isto é, a definição das questões que a ontologia deve ser capaz de responder (*questões de competência*). Estas questões, além de justificar a existência da ontologia, servem para futura avaliação da mesma.

- *Captura da ontologia*

Considerada a fase mais importante, seu objetivo é capturar o conjunto de elementos de

um domínio que podem ser representados em uma ontologia, com base nas questões de competência relacionadas a esta.

A captura envolve a identificação e especificação de conceitos (classes), seus relacionamentos e todos os demais elementos necessários para a representação da ontologia, tais como propriedades, axiomas, instâncias, etc. Falbo [14] sugere ainda que esta especificação seja acompanhada de um vocabulário de termos, que auxiliem no seu entendimento, e seja representada através de uma linguagem gráfica, no sentido de facilitar sua futura comunicação com especialistas de domínio.

Alguns conceitos que não podem ser obtidos a partir de outros conceitos da ontologia (conceitos primitivos) devem ser especificados utilizando-se linguagem natural, com o mínimo de ambigüidade e inconsistência possível. Já os conceitos não primitivos devem ser especificados com referência clara aos conceitos primitivos. A utilização de taxonomias (organização em categoria e sub-categorias) também deve ser utilizada na representação dos conceitos, pois auxilia sua organização, facilitando seu entendimento.

Os conceitos, suas propriedades e relacionamentos formam a base de qualquer ontologia. Entretanto, para definir a semântica de seus termos devem ser construídos os axiomas, que são utilizados para modelar sentenças sempre verdadeiras [2].

Na prática, os axiomas especificam definições de termos na ontologia e restrições sobre sua interpretação. Geralmente especificados em linguagem natural (Falbo [14] cita que os axiomas também podem ser especificados através de lógica de primeira ordem, como sentenças de primeira ordem usando os predicados da ontologia), os axiomas podem ser divididos em dois tipos: *axiomas de derivação* e *axiomas de consolidação*.

Axiomas de derivação são aqueles que permitem explicitar informações a partir do conhecimento previamente existente. Axiomas de consolidação descrevem a coerência das informações existentes, isto é, definem condicionantes para o estabelecimento de uma relação ou para a definição de um objeto como instância de um conceito [14].

Os axiomas de derivação, por sua vez, se dividem em dois tipos [14]: *axiomas epistemológicos* e *axiomas ontológicos*. Os axiomas epistemológicos são descritos para mostrar restrições impostas pela forma de estruturação dos conceitos. Já os ontológicos descrevem restrições de significação impostas no domínio.

Esta classificação dos axiomas é uma boa diretriz para guiar quais são os tipos que devem ser capturados em uma ontologia, ou seja, axiomas que considerem a estruturação dos conceitos e relações (epistemológicos), seus significados e restrições (ontológicos) e as leis de integridade que os regem (consolidação) [14].

O processo de definição de axiomas deve ser fortemente guiado pelas questões de competência. Estes axiomas devem ser necessários e suficientes para expressar estas questões e para caracterizar suas soluções. Além disso, qualquer solução para uma questão de competência deve ser descrita pelos axiomas da ontologia e deve ser consistente com eles.

Caso os axiomas propostos não sejam suficientes para responder às questões de competência, novos axiomas devem ser introduzidos na ontologia, até que esta condição seja verdadeira, assim como axiomas redundantes devem ser eliminados. Desta forma, o processo de captura da ontologia é iterativo e fortemente ligado à avaliação de axiomas.

- *Formalização da ontologia*

Formalização da ontologia corresponde à especificação da ontologia em uma linguagem. Na prática, uma ontologia pode ser representada através de qualquer linguagem não formal (natural), embora a representação formal seja considerada a mais apropriada na maioria dos casos.

Uma linguagem formal é uma linguagem fundamentada em um modelo matemático. Geralmente envolve pressuposições implícitas que entram despercebidas no processo de dedução, tem símbolos não ambíguos e formulações exatas, portanto, a clareza e a correção de uma dedução podem ser testadas com maior facilidade e precisão.

Entretanto, o que determina a escolha da linguagem a ser utilizada na formalização da ontologia depende diretamente do propósito da ontologia a ser desenvolvida. Mais detalhes sobre linguagens formais de especificação estão descritos na Seção 5.3.

- *Integração com ontologias existentes*

De acordo com Falbo [14], durante os processos de captura e/ou formalização, pode surgir a necessidade de integrar a ontologia em questão com outras já existentes, visando aproveitar conceituações previamente estabelecidas. De fato, é uma boa prática desenvolver ontologias funcionais modulares, gerais e que sejam mais amplamente reutilizáveis e, quando necessário, integrá-las, obtendo o resultado desejado.

- *Avaliação*

Avaliar uma ontologia significa verificar se a mesma satisfaz os requisitos definidos em sua construção. Guizzardi [25] sugere que esta etapa seja realizada em paralelo com as fases de captura e formalização. Para isso, podem ser definidos um conjunto de critérios para guiar tanto o desenvolvimento, quanto a avaliação da qualidade das ontologias construídas. Dentre estes critérios estão: clareza, coerência, extensibilidade e compromissos ontológicos mínimos. Além disso, as questões de competência devem ser utilizadas para avaliação dos axiomas.

- *Documentação*

Segundo Guizzardi [25], todo o desenvolvimento da ontologia deve ser documentado. Esta etapa inclui os propósitos, requisitos e cenários de motivação, as descrições textuais da conceituação, a ontologia formal e os critérios de projeto adotados. Assim como a avaliação, esta é uma atividade que deve cobrir todo o ciclo de vida de desenvolvimento da ontologia.

Os termos capturados na conceituação do universo de discurso devem ser descritos em um Dicionário de Termos, considerando dois princípios importantes: o princípio do *vocabulário mínimo* e o princípio da *auto-referência*.

O princípio do *vocabulário mínimo* diz respeito ao vocabulário utilizado na definição dos termos da ontologia. Ele deve ser o menor possível e não deve apresentar ambigüidades. Qualquer termo que não tenha um significado claro e não ambíguo, deve ser formalmente definido no Dicionário.

O princípio da *auto-referência* indica que a definição de um termo no Dicionário deve, sempre que possível, ser feita utilizando outros termos do mesmo.

Com base nestes princípios, o uso de hipertextos surge como uma potencial abordagem para a documentação de ontologias. Esta tecnologia mostra-se adequada, já que torna natural a definição de novos termos a partir de outros mais primitivos, e permite a navegação entre definições, exemplos e a formalização, incluindo os axiomas.

5.2 Ferramentas para Construção de Ontologias

Existem várias ferramentas para construção de ontologias. A seguir são apresentadas três delas. Comparações e/ou mais informações sobre estas e outras ferramentas podem ser obtidas em [2].

- *Protégé 2000* [41]

É um ambiente interativo para construção de ontologias, que oferece uma interface gráfica (Figura 8) para sua edição. Sua arquitetura é modular e permite a inserção de novos recursos. Foi desenvolvido pelo grupo de informática médica da Universidade de Stanford (<http://www.stanford.edu/>).

Possui código aberto em uma aplicação *standalone*, composta por um editor de ontologia e uma biblioteca de *plugins* com funcionalidades. Atualmente importa/exporta para diversas linguagens, dentre elas Flogic [32], OIL [16], XML [9] e Prolog [39].

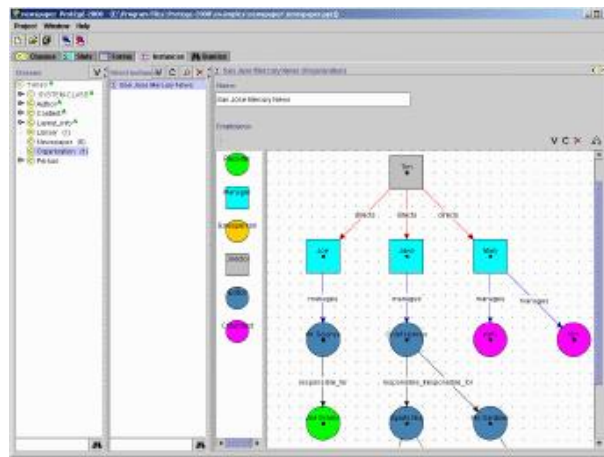


Figura 8: Protégé

- *WebODE* [3]

Foi desenvolvido no laboratório de inteligência artificial da Universidade Politécnica de Madri (<http://www.fi.upm.es/>). É uma aplicação *Web* (Figura 9), onde as ontologias são armazenadas em bases de dados relacionais.

Possui serviços de documentação, avaliação e fusão de ontologias. Dá suporte para Methontology (Seção 5.1). Atualmente importa/exporta, dentre outras linguagens, para XML, RDF(S) [33], OIL, DAML + OIL [28], Flogic e Prolog.

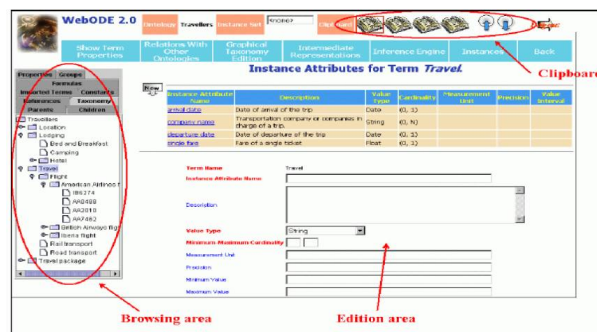


Figura 9: WebODE

- *OntoEdit* [35]

É um ambiente gráfico para edição de ontologias que permite inspeção, navegação, codificação e alteração de ontologias. Foi desenvolvido pela AIFB (*Institut für Angewandte Informatik und Formale Beschreibungsverfahren*) na Universidade de Karlsruhe [31]. Possui arquitetura extensível baseada em *plugins* e atualmente está disponível nas seguintes versões: *OntoEdit Free* e *OntoEdit Professional* (Figura 10).

Possui suporte para *On-to-Knowledge* (Seção 5.1) e importa/exporta para Flogic, XML, RDF(S), DAML+OIL.

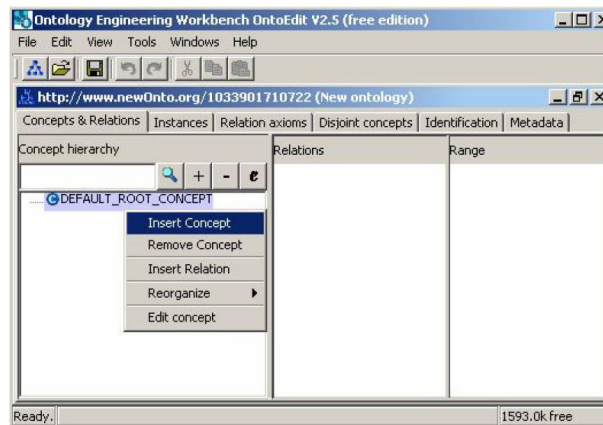


Figura 10: *OntoEdit*

Atualmente, várias ferramentas similares existem disponíveis, mas não há uma ferramenta completa. Seu usuário deve ler sobre as características (descrição, arquitetura, interoperabilidade, representação e usabilidade) de cada uma para escolher aquela que está mais de acordo com os seus objetivos e/ou necessidades.

Não há correspondência entre as metodologias para construir ontologias e ferramentas, exceto para *Methontology* e *WebODE*, assim como *On-To-Knowledge* e *OntoEdit*. Muitas ferramentas apenas focam sobre poucas atividades do ciclo de vida, como projeto e implementação.

Há muitas ferramentas similares disponíveis, porém não são capazes de interoperarem. Isso provoca sérios problemas quando existe a necessidade de se integrar ontologias em bibliotecas de ontologias de diferentes ferramentas ou quando é necessário fazer a fusão de ontologias disponíveis em diferentes ferramentas ou linguagens.

5.3 Linguagens para Representação de Ontologias

As linguagens utilizadas na especificação de ontologias podem ser divididas em três tipos [2]: linguagens de ontologias tradicionais, Linguagens padrão *Web* e linguagens de ontologias *Web-based*.

- *Linguagens de ontologias tradicionais*

Cycl, Ontolíngua, F-Logic, CML, OCML, Loom, KIF.

- *Linguagens padrão Web*

XML, RDF.

- *Linguagens de ontologias Web-based*

OIL, DAML+OIL, SHOE, XOL, OWL.

O critério para escolha de uma linguagem para especificação de ontologias varia de acordo com o tipo a ser especificado. Não é objetivo deste texto tratar este aspecto. Uma comparação entre as principais linguagens pode ser obtida em [48].

Pelo fato de serem as linguagens recomendadas pelo W3C (*World Wide Web Consortium*) [47], a seguir são descritas as principais características da Ontolândia, RDF e OWL. Mais detalhes sobre estas e as outras linguagens citadas nesta Seção podem ser obtidos em [2] e [48].

- *Ontolândia* [10]

Baseada em KIF (*Knowledge Interchange Format*) [18], foi criada em 1992 pelo Laboratório de Sistemas do Conhecimento da Universidade de Stanford. Combina paradigmas de *frames* e predicados de cálculo de primeira ordem. É considerada a mais expressiva de todas as linguagens para representar ontologias.

Permite representar conceitos, taxonomias de conceitos, relações *n-árias*, axiomas, instâncias e procedimentos. Nenhum suporte de raciocínio é fornecido com a linguagem.

Exemplo:

Código 1 – Especificação em Ontolândia

```

1      (define-class female-person (?person)
2          "female humans"
3          :iff-def (and (human ?person)
4                      (= (gender ?person) female)))
```

- *RDF* [36]

Foi desenvolvido pelo W3C como uma linguagem baseada em rede semântica para descrever recursos da *Web*. RDF(S) é a combinação de RDF e *RDF Schema*, que também foi desenvolvido pela W3C como uma extensão do RDF com primitivas baseadas em *frame*.

RDF(S) é bastante expressiva, pois permite a representação de conceitos, taxonomias de conceitos e relações binárias. Algumas máquinas de inferência têm sido criadas para esta linguagem, principalmente para checar restrições.

Exemplo:

Código 2 – Especificação em RDF

```

1      <Class ID="Female">
2          <subClassOf resource="#Animal"/>
3          <disjointWith resource="#Male"/>
4      </Class>
```

- *OWL - Ontology Web Language* [26]

Esta linguagem é recomendada da W3C desde fevereiro de 2004. É uma linguagem para ser utilizada quando as informações contidas em documentos *web*, precisam ser processadas por aplicações em situações em que seu conteúdo precisa mais do que ser apresentado apenas para humanos.

Pode ser usada para representar explicitamente o significado de termos em vocabulários e os relacionamentos entre os termos.

Possui mais facilidades para expressar significados e semântica que XML, RDF e RDF(S), além de representar conteúdo interpretável por máquinas na *Web*. É uma revisão da linguagem DAML + OIL.

Exemplo:

Código 3 – Especificação em OWL

```
1 <owl:AllDifferent>
2   <owl:distinctMembers rdf:parseType="Colecao">
3     <vin:VinhoCor rdf:about="#Tinto" />
4     <vin:VinhoCor rdf:about="#Branco" />
5     <vin:VinhoCor rdf:about="#Rose" />
6   </owl:distinctMembers>
7 </owl:AllDifferent>
8 <owl:AllDifferent>
9   <owl:distinctMembers rdf:parseType="Colecao">
10    <vin:VinhoCorpo rdf:about="#Leve" />
11    <vin:VinhoCorpo rdf:about="#Médio" />
12    <vin:VinhoCorpo rdf:about="#Cheio" />
13  </owl:distinctMembers>
14 </owl:AllDifferent>
15 <owl:AllDifferent>
16   <owl:distinctMembers rdf:parseType="Colecao">
17    <vin:VinhoSabor rdf:about="#Suave" />
18    <vin:VinhoSabor rdf:about="#Moderado" />
19    <vin:VinhoSabor rdf:about="#Forte" />
20  </owl:distinctMembers>
21 </owl:AllDifferent>
22 <owl:AllDifferent>
23 wine.xml
```

Para construir uma ontologia, primeiro é necessário estabelecer o que a aplicação precisa em termos de expressividade. Nem todas as linguagens permitem representar os mesmos componentes e raciocínio do mesmo modo.

Tradutores entre linguagens ainda não são suficientes para assegurar que a informação não será perdida no processo. Por isso, tomar uma boa decisão sobre o uso de uma linguagem específica para representar ontologia é importante para desenvolver aplicações baseadas em ontologias.

Atualmente, uma das linguagens mais utilizadas para representação de ontologias é a XML. Apesar de ser menos expressiva, se comparada a OWL, a linguagem XML é mais popular no ambiente da Internet, seja para a construção de arquivos de configuração, seja para o intercâmbio de dados entre aplicações na *Web* ou estruturação e armazenamento de dados, além de ser suportada pela ferramenta Protégé, que também é utilizada neste trabalho.

Acredita-se que no futuro [9], com a contínua expansão da *Web*, e o advento da *Web-Semântica*, XML se tornará a linguagem universal para representação de dados. Logo, todas as aplicações serão capazes de se comunicar, uma vez que elas poderão entender os vocabulários e/ou marcações de outros documentos produzidos por outras aplicações.

XML (*Extensible Markup Language*) é uma linguagem que permite a construção de documentos legíveis para seres humanos e que podem ser facilmente tratados por máquinas. É um conjunto de regras para a definição de marcadores semânticos, que dividem um documento em partes identificáveis. É uma meta-linguagem que define uma sintaxe para ser utilizada na criação de outras linguagens de marcação para um domínio específico, com estrutura e semântica próprias [27].

O código 4 mostra um exemplo de especificação em XML.

Código 4 – Especificação em XML

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <!DOCTYPE Agenda SYSTEM "C:\Mestrado\Biblioteca.dtd">
3  <bibliografia>
4      <livro>
5          <titulo>
6              Verdes Campinas
7          </titulo>
8          <autor webid="http://exemplo.com/#OP5">
9              Olimpio Pereira
10         </autor>.
11     </livro>
12 </bibliografia>
```

Mais detalhes sobre a linguagem XML podem ser obtidos em [9] e [8].

6 Conclusão

A crescente complexidade dos objetos armazenados na Internet e o grande volume de dados exigem processos de recuperação cada vez mais sofisticados, apresentando novos desafios. De fato, com a explosão do número de documentos e de usuários na *Web*, modelos para recuperação precisa de informações passaram a ser de grande importância.

Em relação as ontologias, dois pontos principais devem ser considerados:

Em primeiro lugar, é necessário trabalhar na criação de ferramentas que facilitem:

- Desenvolvimento de uma ontologia durante todo o ciclo de vida, incluindo: integração, fusão, reengenharia, avaliação de conteúdo, tradução para diferentes linguagens e formatos, troca de conteúdo com outras ferramentas, etc.
- Gerenciamento de ontologias: gerenciamento de configuração e evolução de ontologias isoladas assim como de bibliotecas.
- Suporte de Ontologias: cronograma, documentação, técnicas avançadas para visualização do conteúdo da ontologia, etc.
- Suporte metodológico para construção de ontologias.

Em segundo lugar, é necessário trabalhar na criação dos seguintes serviços:

- Softwares que ajudem a localizar a ontologia mais apropriada para uma dada aplicação;

- Métricas formais que comparam similaridade semântica e distância semântica entre termos da mesma ontologia ou em ontologias diferentes;
- Softwares que permitam incremental, consistente e seletivas atualizações da ontologia, a qual está sendo usada por uma dada aplicação;
- Módulos de *queries* para consultar ontologias;
- Acesso remoto para sistemas de bibliotecas de ontologias;
- Softwares que facilitem a integração da ontologia com sistemas de herança e bases de dados.

7 Agradecimento

Ao Prof. Dr. Cedric Luiz de Carvalho, pela avaliação do presente texto e pelas sugestões feitas, as quais muito contribuíram para a melhoria do texto original.

Referências

- [1] AGUADO, G. E. A. **Ontogeneration: reusing domain and linguistic ontologies for spanish text generation?** In: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, BRIGHTON, ENGLAND, p. 23–28, 1998.
- [2] ALMEIDA, M; BAX, M. **Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção.** Revista Ciência da Informação, 32(3), 2003.
- [3] ARPÍREZ, J. C; CORCHO, O; FERNÁNDEZ-LÓPEZ, M; GÓMEZ-PÉREZ, A. **Webode: a scalable workbench for ontological engineering.** In: INTERNATIONAL CONFERENCE ON KNOWLEDGE CAPTURE. Victoria, British Columbia, Canadá, 2001.
- [4] BERNERS-LEE, T; HENDLER, J; O., L. **The semantic web.** Scientific American, 05 2001.
- [5] BLACKBURN, S; MARCONDES, D. **Dicionário Oxford de Filosofia.** Tradução de Murcho et al., Rio de Janeiro, Jorge Zahar, 1997.
- [6] BORGO, S; GUARINO, N; MASOLO, C; VETERE, G. **Using a large linguistic ontology for internet-based retrieval of object-oriented components.**
- [7] BORST, W. **Construction of Engineering Ontologies for Knowledge Sharing and Reuse.** PhD thesis, University of Twente, P.O. Box 217 - 7500 AE Enschede - The Netherlands, 1997.
- [8] BOSAK, J. **XML, java, and the future of the web.** World Wide Web Journal, 2(4):219–227, 1997.
- [9] BRAY, T; PAOLI, J; SPERBERG-MCQUEEN, C; MALER, E; YERGEAU, F. **Extensible markup language (xml) 1.0.** Disponível em <http://www.w3.org/TR/2004/REC-xml-20040204/>, acessado em abr/2007, 2 2004.

- [10] CHAUDHRI, V. K. **Open knowledge base connectivity 2.0.**
- [11] CLARK, P. **Some ongoing kbs ontology projects and groups.** <http://www.cs.utexas.edu/users/mfkb/related.html>, acessado em novembro de 2006, 2006.
- [12] DOMINGUE, J. **Tadzebao and webonto: discussing, browsing and editing ontologies on the web.** In: 11TH BANFF KNOWLEDGE ACQUISITION WORKSHOP. BANFF, ALBERTA, CANADA., 1998.
- [13] DOMINGUE, J; MOTTA, E. **A knowledge-based news server supporting ontology-driven story enrichment and knowledge retrieval.** In: KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT. DAGSTUHL CASTLE, GERMANY., 1999.
- [14] FALBO, R. A. **Integração de Conhecimento em um Ambiente de Engenharia de Software.** PhD thesis, Universidade Federal do Rio de Janeiro, 1998.
- [15] FELICÍSSIMO, C. H; SILVA, L. F; BREITMAN, K. K; LEITE, J. C. S. **Geração de ontologias subsidiada pela engenharia de requisitos.** In: Luis Eduardo Galvão Martins, X. F, editor, ANAIS DO WER03 - WORKSHOP EM ENGENHARIA DE REQUISITOS, volume 27-28, p. 255–269, Piracicaba-SP, Brasil, 11 2003.
- [16] FENSEL, D; VAN HARMELEN, F; HORROCKS, I; MCGUINNESS, D; PATEL-SCHNEIDER, P. **Oil: An ontology infrastructure for the semantic web.** IEEE Intelligent Systems, 16:38–44, 2001.
- [17] GANDON, F. **Engineering an ontology for a multi-agents corporate memory system.** In: INTERNATIONAL SYMPOSIUM ON THE MANAGEMENT OF INDUSTRIAL AND CORPORATE KNOWLEDGE., p. 209–228, 2001.
- [18] GENESERETH, M. R. **Knowledge interchange format.**
- [19] GERMANN, U. **Making semantic interpretation parser-independent.** In: PROCEEDINGS OF THE AMTA CONFERENCE, 1998.
- [20] GRUBER, T. **Ontolingua: A mechanism to support portable ontologies,** 1992.
- [21] GRUBER, T. **What is an ontology.** <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>, acessado em jan/2005., 1 2005.
- [22] GUARINO, N. **Formal ontology, conceptual analysis and knowledge representation.** International Journal of Human and Computer Studies, 43(5/6):625–640, 1995.
- [23] GUARINO, N. **Understanding, building and using ontologies.** International Journal of Human and Computer Studies, 45(2/3), 2 1997.
- [24] GUARINO, N. **Formal ontology in information systems.** In: IOS Press, A, editor, FOIS'98, TRENTO, ITALY, p. 3–15, 06 1998.
- [25] GUIZZARDI, G. **Desenvolvimento para e com reuso: Um estudo de caso no domínio de vídeo sob demanda.** Master's thesis, Universidade Federal do Espírito Santo, 2000.

- [26] HARMELEN, F. V; MCGUINNESS, D. L. **Owl web ontology language overview.** Disponível em <http://www.w3.org/TR/2004/REC-owl-features-20040210/>, acessado em abr/07, 2 2004.
- [27] HAROLD, E. R. **The XML Bible.** IDG Books, 2 edition, 1999.
- [28] HORROCKS, I; CONNOLLY, D; HARMELEN, F; MCGUINNESS, D; PATEL-SCHNEIDER, P; STEIN, L. **Daml+oil reference description**, 3 2001.
- [29] IDC. **Idc.** <http://www.idc.com>, acessado em julho de 2005, 07 2005.
- [30] JONES, D; BENCH-CAPON, T; VISSER, P. **Methodologies for ontology development.** in Proc. IT KNOWS Conference, XV IFIP World Computer Congress, Budapest, August., 1998.
- [31] KARLSRUHE, U. **Institut für angewandte informatik und formale beschreibungsverfahren.** <http://www.aifb.uni-karlsruhe.de/>, 2006.
- [32] KIFER, M; LAUSEN, G; WU, J. **Logical foundations of object-oriented and frame-based languages.** Journal of ACM, 42:741–843, July 1995.
- [33] LASSILA, O; SWICK, R. **Resource description framework (rdf) model and syntax specification.** W3C Working Draft, p. 1008, 1998.
- [34] LENAT, D. **Cyc: A large-scale investment in knowledge infrastructure.** Communications of the ACM, 38(11):33–38, 1995.
- [35] MAEDCHE, A; STAAB, S; STOJANOVIC, N; STUDER, R; SURE, Y. **Semantic portal - the seal approach**, 2001.
- [36] MANOLA, F; MILLER, E. **Rdf primer.** Disponível em <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>, acessado em Abr/2006, 2 2004.
- [37] MILLER, G. **Wordnet: A on-line lexical database.** International Journal of Lexicography, (3-4), 1990.
- [38] MORENO, A. O; HERNÁNDEZ, C. P. **Reusing the mikrokosmos ontology for concept-based multilingual terminology databases.** INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES E EVALUATION, 2000.
- [39] MUGGLETON, S. **Inductive Logic Programming.** In: THE MIT ENCYCLOPEDIA OF THE COGNITIVE SCIENCES (MITECS). MIT Press, 1999.
- [40] MUSEN, M. A. E. A. **Protege-ii: Computer support for development of intelligent systems from libraries of components.** MEDINFO 95 - WORLD CONGRESS ON MEDICAL INFORMATICS, 8 1995.
- [41] NOY, N. F; MCGUINNESS, D. L. **Ontology development 101: A guide to creating your first ontology.** http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html, Acessado em Nov/2006, 2006.
- [42] ONTOSEARCH. **Welcome to ontosearch2.** <http://www.ontosearch.org/>, acessado em Nov/2006., 11 2006.

- [43] PAUL, Z. Z. **Sharing engineering design knowledge in a distributed environment.**
- [44] SHUM, S. B; MOTTA, E; DOMINGUE, J. **Scholonto: an ontology-based digital library server for research documents and discourse.** Int. J. on Digital Libraries, 3(3):237–248, 2000.
- [45] SMITH, B. **Ontology: Philosophical and computational.** <http://wings.buffalo.edu/philosophy/faculty/smith/articles/ontologies.htm>, acessado em outubro de 2006, 10 2006.
- [46] STAAB, S; SCHNURR, H; STUDER, R; SURE, Y. **Knowledge processes and ontologies.** IEEE Intelligent Systems, 16(1):26–34., 2001.
- [47] W3C. **World wide web consortium.** <http://www.w3.org/>, 2006.
- [48] WACHE, H; OGELE, V; VISSER, T; STUCKENSCHMIDT, U; SCHUSTER, H; NEUMANN, G; UBNER, H. **Ontology-based integration of information - a survey of existing approaches.** citeseer.ist.psu.edu/article/wache01ontologybased.html, 2001.