ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/264156426

Uma revisão da Engenharia do Conhecimento: Evolução, Paradigmas e Aplicações

Article · March 2013

DOI: 10.47916/ijkem-vol2n2-2013-1

CITATIONS
11

READS
3,037

2 authors:

Mara Abel
Universidade Federal do Rio Grande do Sul
186 PUBLICATIONS **1,271** CITATIONS
SEE PROFILE

Sandro Rama Fiorini
IBM
52 PUBLICATIONS **775** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

DIA.BR View project

An Ontological Model for Urinary Profiles View project

All content following this page was uploaded by Mara Abel on 14 April 2016.

The user has requested enhancement of the downloaded file.

International Journal of Knowledge Engineering and Management

UMA REVISÃO DA ENGENHARIA DO CONHECIMENTO: EVOLUÇÃO, PARADIGMAS E APLICAÇÕES

Mara Abel*
Sandro Rama Fiorini**

Resumo

A Engenharia do Conhecimento define metodologias e ferramentas para adquirir e modelar conhecimento com a finalidade de torna-lo independente das pessoas, formalizá-lo e permitir a apropriação por organizações ou sistemas. Desta forma, atende os objetivos da Gestão do Conhecimento. A Engenharia do Conhecimento tem seu foco na criação de modelos formais de um comportamento de solução de problemas contruídos a partir da racionalização da solução observada. Esses modelos são definidos em dois níveis: o nível do conhecimento é voltado a compreensão dos modelos pelas pessoas; o nível simbólico busca a mecanização deste modelo voltado para a eficiência na máquina. O conhecimento, objeto da modelagem, é definido na Ciência da Cognição como um padrão de estados internos que correspondem, cada um, a um estado do universo. A Ciência da Computação argumenta que esses padrões podem ser estados internos de um computador, justificando, assim, que a habilidade inerentemente humana de reter e aplicar conhecimento pode ser exercida pelos computadores. Porém, os estados internos de um computador só encontram sua relação com os estados do universo através da cognição humana. O conhecimento é

* Possui graduação em Geologia pelo Instituto de Geociências (1982), mestrado em Computação pelo Instituto de Informática (1988) e doutorado em Computação pela Universidade Federal do Rio Grande do Sul (2001). É professora associada do Instituto de Informática da Universidade Federal do Rio Grande do Sul. Coordenou o Programa PRIME na UFRGS que permitiu a criação de 98 empresas de base tecnológica. Foi um dos membros do Comitê de Estruturação do Parque Tecnológico da UFRGS e atualmente é membro da Diretoria Executiva do Parque. É coordenador da UFRGS na Rede de Pesquisa em Modelagem de Bacias da PETROBRAS. Fundou a empresa ENDEEPER, como *um spin off* do Grupo de Banco de Dados Inteligentes que coordena. Tem experiência na área de Ciência da Computação, com ênfase em Inteligência Artificial, atuando principalmente nos seguintes temas: Gestão e Engenharia do Conhecimento, desenvolvimento de Ontologias, especialmente aplicados a Geologia de Petróleo – *marabel@inf.ufrgs.br*.

** Possui graduação e mestrado em Ciências da Computação pela Universidade Federal do Rio Grande do Sul (UFRGS). Atuou como engenheiro de software pela empresa ENDEEPER e como professor substituto na UFRGS. Atualmente, é doutorando do Programa de Pós-Graduação em Computação (PPGC) da UFRGS, tendo completado estágio de doutoramento no grupo de Ciências Cognitivas da Universidade de Lund, Suécia. Sua pesquisa tem foco em representação de conhecimento e raciocínio, com ênfase em aspectos ligados ao problema do ancoramento simbólico e abordagens de representação baseadas em espaços conceituais – *srfiorini@inf.ufrgs.br*.

representado por um componente descritivo, que representa as coisas do mundo e um componente dinâmico que extrai novas informações daquelas representadas. Atualmente, a abordagem mais utilizada para representar o componente descritivo são ontologias associadas a modelos de domínio. Ontologias descrevem o significado dos vocabulários compartilhados do domínio, enquanto modelos de conhecimento descrevem as instâncias e particularidades do problema em foco. Por sua vez, o componente dinâmico é descrito através dos métodos de solução de problemas. Ontologias tem se mostrado a abordagem de maior sucesso da Engenharia de Conhecimento. Aplicações de WEB semântica, sistemas especialistas, integração de sistemas, portais de conhecimento atestam a utilidade da representação do conhecimento em modelos formais, que é o objetivo da Engenharia do Conhecimento.

Palavras Chaves: Engenharia do Conhecimento. Gestão do Conhecimento. Definição de conhecimento. Ontologia. Representação de conhecimento.

1 INTRODUÇÃO

A Engenharia de Conhecimento busca capturar o conhecimento das organizações e dos indivíduos e formalizá-lo de maneira independente de domínio de modo a permitir sua apropriação e reuso em sistemas ou processos. Peter Drucker nos permite compreender melhor o papel do conhecimento na citação abaixo:

> O conhecimento que é **agora** considerado conhecimento prova-se a si mesmo somente em **ação**. O quê significa conhecimento é a informação efetiva em ação, a informação focada nos resultados. Os resultados estão **fora** das pessoas, na sociedade e na economia, ou nos próprios avanços do conhecimento. Para realizar qualquer coisa este conhecimento deve ser altamente especializado. Ele não poderia ser aprendido ou ensinado. Nem implica em qualquer princípio genérico. Ele é experiência antes que aprendizado, treinamento antes que estudo. Mas hoje nós não falamos desse conhecimento especializado como **habilidades**. Nós falamos dele como uma **disciplina.** Essa é a grande mudança intelectual jamais registrada. Uma disciplina transforma uma habilidade em uma metodologia - como a engenharia, o método científico, o método quantitativo ou o diagnóstico diferenciado dos médicos. Cada uma dessas metodologias converte uma experiência **ad hoc** em um sistema. Converte anotações em informação. Converte habilidades em algo que pode ser ensinado e aprendido. (DRUCKER, 1993)

Peter Drucker menciona as ideias básicas que norteiam a engenharia de conhecimento hoje:

- o conhecimento não é um bem estático a ser minerado, só existe quando se *produz* algo com ele ou a partir dele;
- o conhecimento não é individual, mas organizacional ou institucional;
- o conhecimento não é genérico, mas associado ou produzido pela solução de uma classe particular de problema ao qual está associado;
- conhecimento não é um conjunto de regras de solução, mas uma experiência que foi sistematizada e pode ser transmitida.

A Engenharia de Conhecimento busca definir metodologias para cobrir as etapas para apropriação do conhecimento, que são: (1) a aquisição do conhecimento de suas fontes humanas ou legadas; (2) a modelagem conceitual, ou a associação dos elementos e conceitos cognitivos a construtos de representação; (3) a representação dos conhecimentos adquiridos através de um ou mais formalismos capazes de capturar a semântica dos conceitos; e (4) sua validação quanto aos objetivos da organização. Essas etapas serão abordadas neste texto.

Em décadas passadas, o desenvolvimento da Engenharia de Conhecimento foi motivado pela incapacidade da indústria de software em reproduzir o sucesso dos seus protótipos ao gerar grandes sistemas especialistas aplicados a problemas reais da indústria. Houve então um esforço com o objetivo de transformar o processo incremental de construir sistemas baseados em conhecimento em uma disciplina da Engenharia baseada em métodos, linguagens e ferramentas especializadas (Studer, Benjamins *et al.*, 1998).

Este artigo busca oferecer um estudo abrangente, porém não aprofundado, da Engenharia de Conhecimento no contexto das aplicações atuais. Ele apresenta a evolução e uma revisão da base conceitual que define a área de estudo, as principais classes de técnicas e ferramentas e, como ilustração, um exemplo de projeto de engenharia de conhecimento para Gestão do Conhecimento, construção de ontologias e projeto de sistemas de conhecimento.

2 A EVOLUÇÃO DA ENGENHARIA DO CONHECIMENTO

As primeiras metodologias da Engenharia de Conhecimento tinham uma abordagem focada na *transferência de conhecimento* dos indivíduos, basicamente os especialistas que detinham o conhecimento, para formalismos de representação e métodos de raciocínio. Os sistemas construídos buscavam reproduzir o mecanismo de solução de problemas dos especialistas nos sistemas gerados, na mesma forma como as pessoas faziam. A compreensão dos métodos de raciocínio individuais mostrou-se um forte limitador desta abordagem, principalmente porque mecanismos humanos de resolução de problemas são fortemente baseados em ponderações de incertezas, além de acoplados e dependentes dos variados modelos mentais que lhe dão suporte. Esses modelos mentais mostraram-se complexos e ineficazes quando transpostos para o computador, sem mostrar contribuição significativa na qualidade da solução gerada pelos sistemas. Sistemas baseados em regras (Buchanan e Shortliffe, 1984), redes semânticas (Duda, Hart *et al.*, 1978) e redes bayesianas (Flores, Hoher *et al.*, 2001) são exemplos que adotaram essa abordagem.

A partir dos anos 90, diversos trabalhos independentes influenciaram os estudos de representação de conhecimento e raciocínio e evoluíam para os novos paradigmas da Engenharia de Conhecimento. Entre esses estudos, os de mais forte influência serão revisados neste artigo: a proposta do Nível do Conhecimento de Alan Newell (Newell, 1982), os estudos sobre ontologias (Gruber, 1995; Guarino, 1998) e os projetos do programa Europeu *Esprit*, números P12 e subsequentes, que buscavam uma metodologia mais estruturada para o desenvolvimento de sistemas baseados em conhecimento, passível de aplicação comercial, cujo resultado deu início a proposição da Metodogia KADS (Schreiber, 1992).

Esses estudos levaram à mudança de paradigma da Engenharia de Conhecimento migrando do enfoque humanístico para uma abordagem de *modelagem*. Nessa visão, o objetivo é criar um modelo da *interação* de um agente inteligente com o meio, de forma a reproduzir os resultados dessa interação em termos de solução de problemas (Velde, 1993). O modelo não necessita inspirar-se em modelos humanos de representação ou inferência, desde que reproduza seu comportamento. A nova visão explora mais a inteligência de máquina do que o

mimetismo da inteligência humana. Desta forma, complexos modelos de raciocínio ou de tratamento de incerteza deixaram de ser o foco principal da pesquisa, agora baseada na construção de modelos robustos, com capacidade de captura semântica e potencialmente reusáveis.

Como reflexo desta mudança, os termos *sistemas especialistas*, que faziam referência à imitação de um especialista humano, e *sistemas baseados em conhecimento*, que eram dirigidos por conhecimento humano, foram substituídos pelo termo mais geral *sistemas de conhecimento* (Schreiber, Akkermans *et al.*, 2000), que define "qualquer sistema de informação que gerencie, armazene e/ou aplique conhecimento organizacional explicitamente representado". O foco atual da Engenharia de Conhecimento é permitir que as organizações se apropriem do seu conhecimento, eventualmente disperso nos trabalhadores de conhecimento, documentos e sistemas, e com ele agreguem valor aos seus produtos e serviços. Desta forma, serve aos objetivos da Gestão do Conhecimento: "criar, adquirir, compartilhar e utilizar ativos de conhecimento"; ou "estabelecer fluxos que garantam a informação necessária no tempo e formato adequados, a fim de auxiliar na geração de ideias, solução de problemas e tomada de decisão" (Boff, 2000).

Desta forma, foram alargados os horizontes de aplicação da Engenharia de Conhecimento. Hoje, a maior demanda de estudos dá suporte a projetos de Gestão de Conhecimento nas organizações (Preece, Flett *et al.*, 2001; Lai, 2007). Outras áreas fortes de aplicação referem-se à construção de Ontologias nos mais variados domínios (Fernandez-Lopez e Gomez-Perez, 2002), especialmente impulsionada pela WEB semântica (Bartalos, Barla *et al.*, 2007). Além destas, a Engenharia de Conhecimento mantém o objetivo original de constituir-se numa metodologia para construção de sistemas baseados em conhecimento (Abel, Silva *et al.*, 2004).

3 PERÍCIA E CONHECIMENTO

Perícia (do inglês *expertise)* é a capacidade de aplicar habilidades intelectuais para resolver problemas em domínios estratégicos, com um desempenho e qualidade de solução superiores à média dos profissionais da área (Abel, 2001). Capturar essas habilidades, estreitamente relacionadas à aplicação do conhecimento e reproduzi-las

através de mecanismos computacionais próprios têm sido o objetivo inicial e maior da Engenharia de Conhecimento.

O termo *conhecimento*, de extenso uso coloquial, foi formalmente definido por estudiosos da Psicologia e da Ciência da Cognição na busca de uma melhor compreensão do comportamento humano. A definição clássica de conhecimento na Epistemologia é "uma crença justificada pela verdade" (Audi, 1998), sem se ater à dificuldade em estabelecer uma *verdade* que é capturada pela mesma mente que define a crença. Essa visão empirista, que considera que todo conhecimento é resultado da experiência, enfatiza conhecimento como um modelo ou uma *versão* adquirida do mundo real, mantida internamente pelo ser humano e eventualmente estendida por mecanismos de abstração próprios da mente (Piaget, 1973).

De fato, o conhecimento é construído na mente que tenta capturar a realidade. Mas essa mente não busca apenas capturar um *espelho* desta realidade. O conhecimento pode ou não corresponder ao padrão externo, pode nem ao menos lembrar os estados externos, ser inato ou aprendido, consciente ou inconsciente. Os estados não representam exatamente a realidade, mas guardam uma interpretação ou uma resposta a ela. Quando aprendem, as pessoas buscam um significado do que estão observando e uma regularidade e ordem nos eventos do mundo que lhes dê sentido (Bodner, 1986). Anderson (Anderson, 1975) enfatizou os aspectos de desenvolvimento interno do conhecimento quando o definiu como "um padrão de estados internos (que pode ser um padrão neural nos animais ou eletrônico em computadores) cada um dos quais correspondendo a algum estado do universo". O conhecimento é o significado construído pela mente, um produto estruturado da reflexão e experiência humana. Faz parte deste conhecimento a capacidade de discernir sua aplicação para lidar com os problemas do mundo. Portanto, somente o ser humano pode ter conhecimento de fato (Kalkan, 2008).

Considera-se que os estados internos de um computador possam também corresponder a uma realidade externa, porém essa será ancorada através da percepção de uma pessoa que restringirá essa realidade e representada em um sistema de símbolos também definido pela cognição humana. Ullman (Ullmann, 1979) demonstra essa correspondência ao afirmar que um sistema de símbolo (em um

computador ou não) representará uma conceitualização na mente de um observador e esses conceitos, por sua vez, terão uma correspondência com a realidade (Figura 1). Não há relação entre uma representação em computador e a realidade sem um filtro da cognição humana.

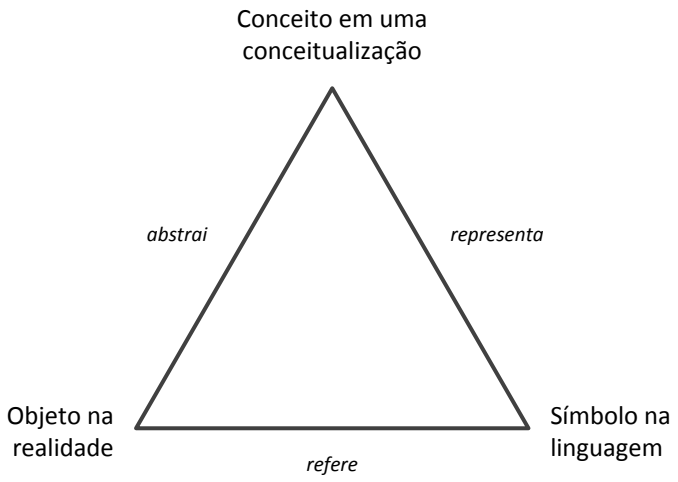

Figura 1 – O Triângulo de Ullman apresenta a relação entre um objeto e seu conceito na mente do observador e uma possível representação em uma linguagem (Ullmann, 1979).

No contexto da Ciência da Computação, conhecimento foi definido por Hayes-Roth e colegas como "descrições simbólicas que caracterizam os relacionamentos empíricos e definicionais num domínio e os procedimentos para manipulação dessas descrições" (Hayes-Roth, Waterman *et al.*, 1983).

Essa definição de conhecimento enfatiza a dualidade do conceito que refere a:

- um componente *descritivo* que se ocupa em representar os objetos (ou as *coisas*) do mundo, seu significado e como esses objetos se relacionam entre si;
- um componente *dinâmico* que *sabe* como essas descrições podem ser utilizadas.

O componente dinâmico do conhecimento inclui os procedimentos de manipulação de descrições e também a capacidade de reconhecimento dos problemas do mundo e as possíveis estratégias que podem ser aplicadas em sua solução.

No âmbito da Engenharia do Conhecimento, o componente descritivo é normalmente associado a uma representação simbólica, através de algum formalismo de representação, cuja expressão mais atual são as linguagens ontológicas. Em termos cognitivos, porém, grande parte dos objetos do mundo são componentes sensórios do conhecimento, dificilmente traduzíveis por representações simbólicas. Como demonstrado em (Polanyi, 1966; Abel, Mastella *et al.*, 2004; Kosslyn, 1994; Hoffman, 1987; Barry, 1997; Burton, Bruce *et al.*, 1999) embora os objetos sensórios sejam essenciais à tomada de decisão, muitas vezes seu uso é inconsciente (Polanyi, 1966; Nonaka e Takeuchi, 1997), dirigido pela atenção atribuída ao problema (Santin, 2007) e de difícil aquisição e representação (Abel, Mastella *et al.*, 2004; Santin, 2007).

Uma definição satisfatória do termo conhecimento, tanto do ponto de vista cognitivo como de Inteligência Artificial, afirma que conhecimento deve ser compreendido como "a informação sobre o domínio e a forma como essa informação é utilizada para resolver problemas" (Abel, 2001). Compreende-se aqui que a *informação* inclui os componentes conceituais do conhecimento, mas também imagens, sons, padrões perceptivos e outros modelos mentais. Os aspectos dinâmicos do conhecimento referem-se ao modo de utilização desse conhecimento para resolver problemas. Todos esses componentes do conhecimento são ainda objeto de intenso estudo na Engenharia de Conhecimento

Para compreender os formalismos de representação de conhecimento, é útil classificar os tipos de conhecimento. Michael Polanyi (Polanyi, 1966) propõem a divisão entre *conhecimento tácito* (ou implícito) *e explícito*. Nonaka e Takeuchi aplicaram esses conceitos no contexto da Gestão do Conhecimento em (Nonaka e Takeuchi, 1997), afirmando que o conhecimento tácito é pessoal, sensorial, específico ao contexto e difícil de ser formulado e comunicado. O conhecimento explícito que pode ser expresso em palavras e números representa apenas uma parcela do conjunto do conhecimento como um todo.

O conhecimento pode ser também classificado como:

- *declarativo*: descreve o que são as coisas do mundo, na forma de *conceitos e suas propriedades*, e como essas coisas se organizam ou

interferem umas nas outras, descrito como *relações entre conceito*. É o foco de representação das ontologias;

- *inferencial*: componente dinâmico do conhecimento que diz de que forma os conceitos podem ser manipulados por mecanismos de inferência para extrair novos conhecimentos úteis;
- *semântico*: expressa o significado dos conceitos e também sua equivalência.

Por exemplo, faz parte do conhecimento declarativo afirmar que *doença* tem *causa*, *sintomas* e provoca risco de vida a um *paciente*. O conhecimento inferencial explora a relação entre sintomas e doenças para extrair a causa das doenças, aplicando um raciocínio do tipo hipótese e verificação. O conhecimento semântico expressa o significado dos conceitos. Nesse exemplo, definir que *doença* é um mal que acomete o homem e que tem o mesmo significado de *moléstia*.

4 A NOÇÃO DE NÍVEL DE CONHECIMENTO DE NEWELL

A noção de nível do conhecimento, introduzida por Allen Newell nos anos 80 (Newell, 1982), tem proporcionado uma perspectiva comum para pesquisadores da Engenharia do Conhecimento e impulsionou os resultados dos projetos de construção de bases de conhecimento. Newell propôs que o problema de capturar conhecimento deveria ser tratado em dois níveis distintos, separando os aspectos semânticos e de compreensão de domínio dos problemas de implementação de sistemas. Nesta abordagem, desenvolver um sistema de conhecimento é visto como a construção de um conjunto de modelos relacionados a algum comportamento de solução de problemas e propor a eles uma tradução computacional.

O domínio é analisado como se um agente (que pode ser o especialista ou o que poderá vir a ser o sistema de conhecimento) possuísse conhecimento sobre o mundo e utilizasse esse conhecimento de modo completamente racional para atingir seus objetivos (Velde, 1993). O nível do conhecimento permite a descrição do *comportamento* em termos de qual conhecimento que ele utiliza (*Modelo do Quê*), como ele utiliza (*Modelo do Como*) e para qual finalidade (*Modelo do Porquê*). Esses modelos descrevem os conceitos do domínio, as inferências possíveis e as

utilizadas pelo agente e a tarefa que esta sendo resolvida. Já o nível simbólico fornece a descrição do mecanismo que permite reproduzir esse comportamento e que atua sobre símbolos e estruturas de símbolos. O nível simbólico é *orientado a sistema,* enquanto o nível do conhecimento é *orientado para o domínio*. Um modelo no nível do conhecimento representa o conhecimento que *racionaliza* aquele comportamento.

É importante notar que o objeto da modelagem no nível do conhecimento não é conhecimento, mas sim comportamento, ou seja, a interação observada entre um agente e seu ambiente (Clancey, 1989). Essa noção, de fato, é que provocou a mudança de paradigma na Engenharia de Conhecimento. Um modelo no nível de conhecimento é um modelo de comportamento nos termos do conhecimento, exatamente como um modelo no nível simbólico é um modelo de interação em termos de símbolos e representações (Figura 2). O que mantém esses modelos juntos é o fato de que eles modelam a mesma coisa, ou seja, a interação observada. Os modelos no nível do conhecimento e no nível simbólico são coerentes e consistentes na extensão do que se propõem a modelar e, em certo grau, devem levar à mesma classe de comportamentos.

A evolução das noções apresentadas por Newell levou ao surgimento de uma série de metodologias de aquisição e representação de conhecimento. Essas metodologias tiveram como maior foco os problemas de aquisição e modelagem de conhecimento, como as Tarefas genéricas (Chandrasekaram, 1986); Métodos de limitação de papéis (Mcdermott, 1988); Componentes da perícia (Steels, 1990); Ontologias (Wielinga e Schreiber, 1994) e OntoLíngua (Gruber, 1992); VITAL (Stutt e Motta, 1994), PROTÉGÉ (Gennari, Musen *et al.*, 2003) e ONTOKEM (Rautenberg, Gomes Filho *et al.*, 2010) Já as metodologias MOKA (Stokes, 2001) e CommonKADS (Schreiber, Akkermans *et al.*, 2000) foram concebidas não apenas para aquisição e modelagem de conhecimento, mas para cobrir todo o ciclo de vida de um projeto de sistema de conhecimento. CommonKADS, apesar das lacunas da metodologia em termos de engenharia de software (Castro, Victoreti *et al.*, 2008) tornou-se o padrão ainda adotado na indústria para o desenvolvimento de sistemas de conhecimento.

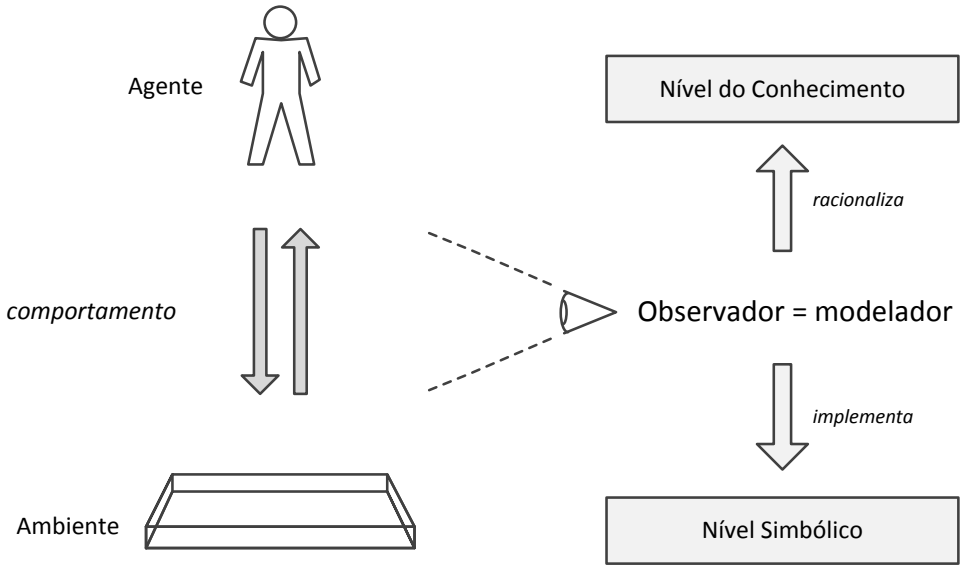

Figura 2 – O nível do conhecimento e o nível simbólico são modelos do comportamento, ou seja, da interação observada entre o agente e o ambiente. Adaptado de (Velde, 1993).

5 ONTOLOGIAS

O termo *ontologia* se refere a um ramo da Filosofia dedicado ao estudo da existência. As noções filosóficas da existência do ser se mostraram adequadas ao objetivo de modelar abstratamente as *coisas* do mundo. Gruber (Gruber, 1993) introduziu o conceito que foi posteriormente refinado por Borst (Borst, 1997) para o contexto da Ciência da Computação: "Ontologia é uma especificação formal de uma conceitualização compartilhada."

Nicholas Guarino em (Guarino, 1998) analisa essa definição enfatizando a importância do compromisso entre a representação e a intenção do modelador e de que forma esse compromisso restringe as possibilidades de representação da própria linguagem: **"**Ontologia é uma teoria lógica que lida com o significado pretendido de um vocabulário formal, i.e., seu compromisso ontológico com uma conceitualização particular do mundo."

Ontologia une os requisitos da modelagem no nível do conhecimento, ao se referir a conceitos comuns a um grupo de pessoas que compartilham seu significado, e também no nível simbólico ao afirmar que esses conceitos devem ser processáveis por computador. Os dois requisitos são comumente antagônicos: a maior

expressividade semântica exige construtos de representação mais ricos e, portanto, mas difíceis de serem processados sem ambiguidade. O equilíbrio desses dois requisitos orienta a pesquisa na área de ontologias atualmente a dois objetivos: a definição de um conjunto de primitivas de representação necessárias e suficientes para modelar as conceitualizações; e a proposição de novas linguagens capazes de expressar a semântica dessas primitivas de forma eficiente. O estudo de conceitualizações e sua expressão formal, base para a construção de modelos de conhecimento e ontologias tem sido foco de intensivo estudo por Nicholas Guarino e colegas (Guarino, 1995; Guarino e Welty, 2004; Guizzardi, 2005). Para a área da Ciência da Computação, a abordagem de ontologias trouxe uma solução para a necessidade da formalização dos modelos da porção do conhecimento que é compartilhável (notadamente, o vocabulário especializado) em formatos e conteúdos independentes da aplicação, permitindo assim seu compartilhamento e reusabilidade.

Como ontologias podem ser utilizadas para estudos de modelagem conceitual e também para construir bases de conhecimento, Gómez-Perez e colegas em (Gómez-Pérez, Fernández-López *et al.*, 2007) propõem uma separação de ontologias em dois tipos úteis. De um lado as *ontologias de representação*, que descrevem primitivas de modelagem capazes de representar o significado dos conceitos. Por outro lado, as *ontologias de domínio* formalizam conceitos restritos a um determinado domínio de aplicação com o objetivo de compartilhar seus significados.

Ontologias de domínio diferem de modelos de conhecimento mais por sua finalidade e restrições do que por sua estrutura interna. Um modelo de conhecimento se propõe a capturar o conhecimento necessário à solução de um problema ou de uma classe restrita de problemas. Com o foco neste objetivo, o conhecimento é adquirido de uma ou poucas pessoas que o dominam, normalmente um especialista no problema em questão. A formalização, onde conceitos abstratos e instâncias do conhecimento se misturam, tem como foco a implementação que favoreça o processamento do raciocínio. Desta forma modelos de conhecimento são construídos voltados ao método de solução de problemas eliciado do especialista e, portanto, são fortemente conectados ao sistema de conhecimento que os executam.

Ontologias de domínio, por sua vez, têm por objetivo a captura de um conhecimento consensual, não individual, para que seja formalizado e compartilhado em uma comunidade de interesse, minimizando as ambiguidades do domínio. São fortemente independentes de implementação e suas linguagens buscam um formalismo rico o suficiente para expressar a semântica dos conceitos, mas capaz de ser processado e reutilizado por muitos tipos diferentes de sistemas.

Muitas ontologias de domínio têm sido construídas com o objetivo inicial de dar suporte a métodos de solução de problemas e também para oferecer um vocabulário comum para o processamento de informações na WEB. Ontologias de domínio e métodos de solução de problemas representam respectivamente os componentes declarativo e inferencial do conhecimento e são abordagens complementares na construção de sistemas de conhecimento (Gómez-Pérez e Benjamins, 1999).

6 ETAPAS NA CONSTRUÇÃO DE BASES DE CONHECIMENTO

Embora a Engenharia de Conhecimento se preocupe com todos os aspectos do desenvolvimento de um sistema de conhecimento, a construção de bases de conhecimento mostra-se o maior desafio. Uma base de conhecimento materializa a apropriação do conhecimento de solução de problemas de uma organização e, portanto, tem valor *per si*.

Bases de conhecimento podem ser desenvolvidas para formalizar o conhecimento e capturá-lo, permitir o compartilhamento de experiências e habilidades em uma organização ou dar suporte ao raciocínio de sistemas especialistas.

A construção de uma base de conhecimento se dá através de três etapas: a aquisição do conhecimento, a modelagem conceitual e a representação do conhecimento. Essas etapas não são estanques e independentes, mas desenvolvem-se na forma de uma espiral de resultados crescentes. A aquisição se refere à compreensão do domínio e da cognição, a segunda implica em escolhas de visões e porções do conhecimento a ser capturado, enquanto a terceira se refere à conexão dos modelos no nível do conhecimento para aqueles do nível simbólico.

a. Aquisição de conhecimento

A aquisição de conhecimento é o processo de extrair, estruturar e organizar o conhecimento de uma ou mais fontes. É a atividade inicial do processo de engenharia de conhecimento, realizada por um engenheiro de conhecimento, que é o agente que exerce o duplo papel de compreender o domínio da informação e de ser capaz de fazer julgamentos em relação às limitações de implementação do sistema. É a tarefa mais crítica de um projeto de Gestão de Conhecimento, pois ainda não existe uma metodologia confiável de aplicação que resulte em um modelo do domínio e dos processos cognitivos de seus agentes. Mesmo assim, é uma tarefa indispensável para que as informações essenciais sejam coletadas e o conhecimento-chave fique disponível para ser organizado, representado, implementado e validado através de um sistema.

Mastella em (Mastella, 2004) separa as técnicas de aquisição de conhecimento de acordo com as fontes do conhecimento, sejam elas não-documentadas ou documentadas (livros, documentos, mapas, diagramas, sistemas, filmes); de acordo com o uso ou não do suporte do computador ao processo de aquisição (Boose, 1991); e de acordo com o tipo do conhecimento a ser adquirido. O termo *eliciação* é utilizando quando o conhecimento é não-documentado e deve ser adquirido de fontes humanas, basicamente através de entrevistas, análise de protocolos (registros de discursos, diálogos, etc) ou com suporte de computador (Cooke 1994). A aquisição de conhecimento de fontes documentadas pode ser feita utilizando diferentes abordagens, como análise manual de documentos, ou técnicas semi-automáticas ou automáticas, como algoritmos de extração de conteúdo de documentos, mineração de dados, raciocínio baseado em casos, aprendizado de máquina, entre outras. Cada uma dessas abordagens constitui uma área de pesquisa em si, enquanto a eliciação de conhecimento é tópico de estudo inserido na Engenharia de Conhecimento e da Ciência da Cognição.

b. Modelagem conceitual

O resultado do estudo do domínio e da aplicação das técnicas de aquisição de conhecimento é um conjunto de conceitos, relacionamentos, regras, procedimentos e

métodos, além de requisitos de sistemas. Essas informações geram relatórios não estruturados que irão compor a documentação total do projeto. Os conceitos, relacionamentos, regras e métodos, por sua vez darão origem ao modelo conceitual, no nível do conhecimento, que será posteriormente formalizado através de uma ou mais abordagens de representação, no nível simbólico. A modelagem conceitual se refere, portanto, a captura semântica do domínio, ou a construção da *conceitualização* ontológica.

A construção do modelo conceitual é a etapa-chave da construção de um modelo de conhecimento ou de uma ontologia de domínio. O engenheiro de conhecimento irá definir quais porções do domínio devem ser modeladas e quais permanecem para etapas futuras ou apenas como documentação. Os conhecimentos selecionados serão então *aproximados* às primitivas (ou construtos) de representação que mais expressam seu significado.

As primitivas de representação são construtos semânticos universais ou semi-universais a todas as linguagens de representação e devem possuir uma relação única com os elementos da sintaxe das linguagens. Os principais construtos utilizados para modelagem conceitual (Gruber, 1995; Guarino, 1995) podem ser resumidos por:

$$\mathrm{MC} = \langle \mathcal{C}, \mathcal{R}, \mathcal{F}, \mathcal{A}, \mathcal{I} \rangle,$$

onde $\mathcal{C}$ descreve o conjunto de *classes* ou *conceitos* que representam elementos de um domínio e que preservam uma relação 1:1 com objetos a serem modelados. Seja P um conjunto de atributos e D um conjunto de codomínios de valores que um atributo pode assumir em um dado modelo, é possível definir que para cada $C \in \mathrm{C}$, tem-se que $C = \{(p, d) \mid p \in \mathrm{P}, d \in \mathrm{D}\}$. R é um conjunto de relações que representam tipos de associações entre elementos do domínio. $\mathcal{F}$ é um conjunto de relações especiais que mapeiam um ou mais elementos do domínio para um único elemento e para as quais podemos atribuir semântica específica. Funções podem modelar regras para dar suporte a inferência. O conjunto A agrupa axiomas que permitem definir declarações sobre os demais elementos do modelo que restringem o domínio. Por fim, I é um conjunto de instâncias que representam os indivíduos do domínio a ser modelado, normalmente mantendo uma relação 1:1 com os objetos existentes no mundo real, sejam eles concretos ou não.

Os relacionamentos podem ser estruturais (que auxiliam na organização do próprio domínio) ou eventuais, que descrevem relacionamentos entre algumas classes apenas. Os relacionamentos estruturais mais importantes são os *taxonômicos* (classe-subclasse) e os *mereológicos* ou *partonômicos* (classe-componente) (Fernández-Breis e Martínez-Béjar, 2002). Associações são relacionamentos que podem ter semântica definida pelo modelador.

Além dessas primitivas, diferentes autores estudam novas primitivas de modelagem de conhecimento. Mastella em (Mastella, Abel *et al.*, 2005) estendeu o conjunto de primitivas para modelar características espaço temporais do domínio, definindo, incluindo primitivas tais como *eventos*, para identificar acontecimentos que alteram o domínio; *relações temporais*, que descrevem as relações de ordem entre os eventos (*antes, depois, durante*, etc); e *relações espaciais*, que descrevem as relações de posição espacial entre as classes do modelo.

A construção de um modelo capaz de capturar conhecimento baseia-se fortemente na escolha correta dos construtos de modelagem capazes de expressar o conhecimento do domínio e perfeita compreensão do seu significado. A intensa pesquisa de novas primitivas de modelagem e sua tradução para linguagens de representação demonstram as dificuldades em expressar objetos cognitivos de maneira formal. Os maiores avanços têm acontecido no estudo e proposição de *Ontologias de Fundamentação* que tem inovado em duas frentes:

(1) através da especialização das primitivas de modelagem que tradicionalmente tem definido as linguagens de representação;

(2) pela restrição das possibilidades de ancoramento simbólico dessas primitivas aos objetos do mundo, ao definir as propriedades essenciais que esses objetos devem possuir para serem por ela representados.

Nessa linha, são notáveis os trabalhos de Guarino e Welty (Guarino e Welty, 2002; Guarino e Welty, 2004) ao definirem a metodologia ONTOCLEAN[1]. A metodologia oferece um guia para a escolha das primitivas de modelagem com base em noções filosóficas de análise metafísica e que, por sua vez, definem

[1] http://www.ontoclean.org/

metapropriedades que impõem restrições para a construção dos modelos. Os autores definem as metapropriedades de essência, rigidez, identidade e unidade para a análise dos conceitos e de suas representações, desta forma, auxiliando na separação da noção de objetos concretos (*estátua*) e de sua substância (*cerâmica*), de objetos com identidade própria (*pessoa*) de outros que tem sua identidade criada por outros objetos (um *estudante* é um conceito que depende da existência de uma *escola* e de uma *pessoa* para existir).

As metapropriedades definidas por Guarino e Welty propiciaram o aperfeiçoamento das linguagens de ontologias, ao permitir a especialização das primitivas de representação em construtos mais precisos e de significado mais restritos. A *Unified Foundational Ontology* (UFO) definida por Giancarlo Guizzardi (Guizzardi, 2005) cria uma hierarquia de metaconstrutos de representação de objetos, relações e propriedades especializando esses construtos a partir de noções filosóficas de propriedades ontológicas. A UFO(A) define os construtos para a representação de objetos que persistem no mundo (continuantes), enquanto a UFO(B) permite a representação de objetos temporais (ocorrentes), tais como eventos, participantes e relações temporais, e a UFO(C) representa os objetos sociais e suas relações (por exemplo, organizações e contratos). Mesmo não dotada de linguagem de representação passível de compilação, a UFO tem influenciado um grande numero de trabalhos de pesquisa em ontologias no Brasil.

A Figura 3 apresenta uma representação diagramática de um modelo conceitual parcial de uma loja de ferragen baseado na UFO(A). Os retângulos rotulados com *kind* representam objetos com identidade ontológica enquanto aqueles com *role* indicam conceitos cujas instâncias são instâncias de algum *kind*. Enquanto os arcos representam relações, os retângulos rotulados com *relator* materializam os conceitos que dão significado e aridade a essas relações. Um modelo conceitual presume a especificação completa do conjunto dos atributos e domínios dos conceitos, portanto o conjunto de duplas atributo/domínio para cada um dos conceitos, não representado no diagrama, deve ser descrito como parte do modelo. Um modelo conceitual, por outro lado, não é dependente da escolha de sua

apresentação. Utilizar representações diagramáticas ou textuais, ou mesmo ambas, depende da conveniência do modelador.

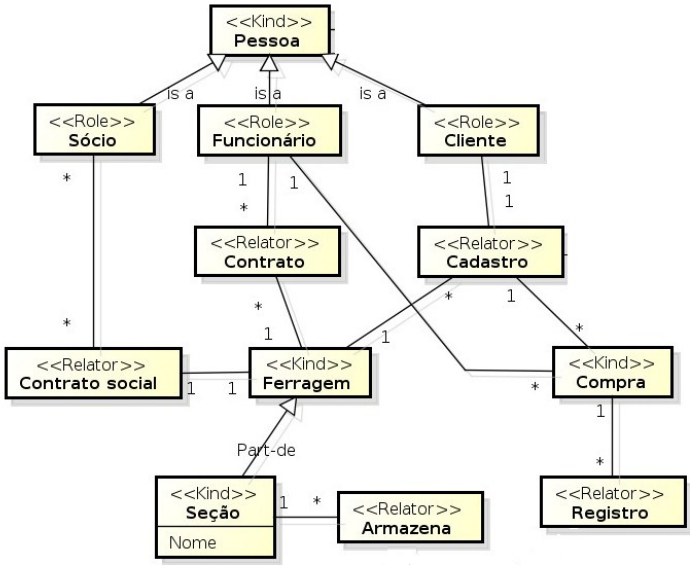

Figura 3 – Representação diagramática de uma parte de modelo conceitual para uma ferragem baseada na UFO(A).

Modelos conceituais não têm por objetivo apenas a descrição abstrata do conhecimento declarativo. Novas abordagens buscam a representação do conhecimento inferencial com modelos conceituais no nível do conhecimento. Entre as mais promissoras estão as estruturas de inferência de Common KADS (Schreiber, Akkermans et al. 2000), os padrões cognitivos de Gardner (Gardner, 1998) e os Métodos de Solução de Problemas (Gómez-Pérez e Benjamins 1999; (Fensel e Motta, 2001)}. As três abordagens têm em comum a proposição de uma representação diagramática de uma sequência abstrata de passos de inferência sobre conceitos do domínio. Métodos de Solução de Problemas atingiram uma maior especificação nos modelos ao incluir na definição dos métodos a competência (o quê o método pode realizar); a descrição operacional do método em termos dos papéis do conhecimento associados aos conceitos, as ações de inferência e a especificação de cont*role*; e os *requisitos e suposições* de aplicação do método. Cada uma das abordagens busca definir um padrão de raciocínio específico a uma classe de problemas, porem genérico em relação a reusabilidade em diferentes domínios. Silva

em (Silva, 2001) detalha a definição e aplicação de um método de solução de problemas para interpretação de ambientes diagenéticos de rochas siliciclásticas.

c. Representação do Conhecimento

Embora modelos conceituais possam ser criados para diferentes finalidades, o processamento por computador requer que o conhecimento seja descrito em uma representação *formal*, ou seja, uma representação não-ambígua e processável por computador. Para isso, existem as linguagens de representação de conhecimento. Elas disponibilizam primitivas de representação que buscam capturar e estruturar conceitos de um determinado domínio, ao mesmo tempo em que retém a sua representatividade semântica.

As principais linguagens de representação de conhecimento são baseadas em dois paradigmas dominantes. O paradigma de *frames*, proposto originalmente por Marvin Minsky (Minsky, 1975; Brachman e Levesque, 2004) se assemelha em parte à modelagem orientada a objetos. Nesse paradigma, as entidades do domínio são representadas na forma de *conceitos, propriedades, restrições e instâncias*, além de incluir conceitos como herança e *demons* sobre propriedades. Já no paradigma de Lógica de Descrição (DL, do inglês *Description Logics*) (Baader, Calvanese *et al.*, 2004), o conhecimento é capturado e formalizado como axiomas lógicos, onde conceitos são representados por suas propriedades necessárias e suficientes. Em geral, DL é mais flexível, pois permite combinar conceitos em construções mais sofisticadas e de maior valor semântico. Uma das principais características da DL são as suas propriedades computacionais (completude, complexidade, computabilidade) já bastante exploradas e delimitadas pela literatura científica. Linguagens baseadas no paradigma DL exploram essas características a fim de apresentar um arcabouço de modelagem e inferência mais consistente e robusto.

Entre a década de 80 e o início da década de 90 diversas linguagens de representação foram desenvolvidas tendo como base os paradigmas de frames e DL. As linguagens CycL (Lenat e Guha, 1990), Ontolingua (Farquhar, Fikes *et al.*, 1997) e FLogic (Kifer, Lausen *et al.*, 1995) se basearam em uma combinação de frames com lógica de primeira ordem. Já as linguagens LOOM (Macgregor, 1991)e KL-

ONE (Brachman e Schmolze, 1985) utilizavam o paradigma de DL. Embora essas linguagens atendessem o problema de representação semântica, a sintaxe ainda era um desafio. Por exemplo, sistemas de conhecimento de maior porte eventualmente exigiam que se acoplassem subsistemas usando modelos formalizados em linguagens distintas. O desafio de integrar bases de conhecimento heterogêneas levou ao desenvolvimento de iniciativas como o protocolo OKBC (K.Chaudhri, Farquhar *et al.*, 1998). O protocolo OKBC (do inglês, *Open Knowledge Base Connectivity*) buscava fornecer um meio eficiente de integração de bases de conhecimento heterogêneas buscando a integração sintática e semântica dos seus formalismos de representação. Por exemplo, as linguagens LOOM, CycL e Ontolingua eram compatíveis com o protocolo OKBC (Gómez-Perez, 2003).

Juntamente com o surgimento da Web surgiram também as linguagens de marcação, como HTML (Raggett, Hors *et al.*, 1999) e XML (Bray, J. Paoli *et al.*, 2006). Essas linguagens representavam uma padronização na sintaxe para representação de informações através dos diversos repositórios de dados existentes na rede. Por outro lado, a linguagem e a marcação careciam de uma representação que agregasse semântica aos dados. Essa necessidade deu origem ao conceito de *Web Semântica* (Berners-Lee, Hendler *et al.*, 2001) e à segunda geração de linguagens para representação de conhecimento. Essas linguagens foram definidas com base em sintaxes padrão (predominantemente XML) e, por isso, resolviam em grande parte o problema de integração existente nas linguagens de primeira geração. A primeira das novas linguagens a ter um maior destaque foi o RDF (Beckett, 2004), introduzido pela W3C (*World Wide Web Consortium*) em 1999. O RDF é equivalente ao paradigma de redes semânticas, o que lhe garante bastante flexibilidade de representação. Em 2002, a W3C propôs a uma extensão de RDF que abarcava também primitivas do paradigma de frames, denominada RDF *Schema* (Brickley e Guha, 2004). Essas primitivas permitem a representação explicita de classes, propriedades e restrições. A junção de RDF e RDF Schema é conhecida como RDF(S).

Três outras linguagens foram derivaram de RDF(S): OIL, DAML+OIL e OWL. A linguagem OIL (Fensel e Motta, 2001) estendeu a capacidade de expressão

de RDF(S) com outras primitivas de frames e uma semântica baseada em DL. DAML-OIL (Horrocks, 2002) incrementou a especificação de OIL com outras primitivas de representação de conhecimento DL. As experiências com o desenvolvimento destas duas linguagens levou a criação da linguagem OWL (Bechhofer, Harmelen *et al.*, 2004), recomendada pela W3C como uma extensão de RDF(S) para representação de ontologias na Web. Atualmente, OWL é uma das linguagens de representação de ontologias mais populares. Contribui para isso a existência de ferramentas livres para criação de ontologias[2], suporte para persistência em banco de dados[3] e pacotes para manipulação de OWL por linguagens de programação[4],

Grande parte do esforço por trás do desenvolvimento do OWL foi preservar algumas propriedades computacionais da DL, como completude computacional e decidibilidade – desejáveis em mecanismos de classificação –, sem que se perdesse muito da expressividade semântica da linguagem. A especificação da W3C define três subconjuntos da linguagem OWL baseado na sua capacidade de representação e propriedades formais: *OWL Lite*, *OWL-DL* e *OWL Full*. OWL Lite é o subconjunto com menor expressividade. Possui construtos para representação de taxonomias simples e algumas restrições sobre propriedades. O OWL-DL estende ao máximo a expressividade da OWL Lite, mantendo as suas propriedades computacionais de complexidade e decidibilidade. O conjunto mais completo, OWL Full, permite o maior nível de representatividade de OWL, permitindo construções de meta-modelagem. No entanto, em OWL Full não há garantias computacionais para máquinas de inferência.

Algumas linguagens de representação possuem primitivas para representação também de conhecimento inferencial. A linguagem CML, por exemplo, que faz parte da metodologia CommonKADS (Schreiber, Akkermans *et al.*, 2000), dispõe de um arcabouço para representação de conhecimento declarativo e também de raciocínio. Enquanto o componente declarativo é representado em um modelo baseado em

---

[2] Protégé Ontology Editor (http://protege.stanford.edu/)
[3] Jena (http://jena.sourceforge.net/)
[4] OWL API (http://owlapi.sourceforge.net/) e LinqToRdf (http://code.google.com/p/linqtordf/)

frames, o conhecimento inferencial é modelado como métodos de raciocínio genéricos e reutilizáveis, com uma linguagem procedural de alto nível.

Atualmente, a linguagem OWL tem sido a mais utilizada para aplicações da WEB Semântica e para formalização de ontologias, embora tenha construtos ainda limitados para a representação do significado dos conceitos. A consolidação de ferramentas como PROTÉGÉ, que auxilia o processo de construção e formalização dos modelos, tem auxiliado grandemente na sua disseminação. CommonKADS e suas ferramentas, por sua vez, têm sido adotadas no meio empresarial como suporte a engenharia de software de sistemas de conhecimento.

Paralelamente, os estudos sobre Ontologias de Fundamentação têm requerido linguagens mais expressivas e com ancoramento simbólico mais restrito do que as acima descritas. O desafio do estado da arte em representação de conhecimento é a definição de linguagens ontológicas mais expressivas e de maior granularidade nos seus construtos de representação, porem que garantam a computabilidade e decidibilidade da DL e suas linguagens derivadas.

7 APLICAÇÕES DA ENGENHARIA DE CONHECIMENTO

A engenharia de conhecimento vem sendo aplicada na construção de projetos de Gestão de Conhecimento, sistemas de conhecimento, e aplicações de processamento e busca de informações na WEB.

A Gestão de Conhecimento demanda métodos de aquisição de conhecimento para identificar e capturar o conhecimento em suas diferentes fontes. Os modelos conceituais e formalismos de representação, mesmo quando estão desvinculados de projetos de sistemas de informação e de conhecimento, oferecem opções para documentação do conhecimento e sua apropriação.

Por sua vez, a construção de sistemas que apliquem conhecimento humano na solução de problemas que envolvem alta perícia ainda requer técnicas e ferramentas mais adequadas para a eliciação, modelagem e formalização do conhecimento do que as disponíveis atualmente e continua impulsionando a pesquisa na área. A mudança de abordagem da Engenharia do Conhecimento, acontecida nos anos 90, permite hoje um melhor domínio no processo de engenharia destes sistemas, quando o

conhecimento envolvido não requer tratamentos de incertezas sofisticados, como aqueles utilizados pelos sistemas das décadas de 70 e 80. Modelos de conhecimento mais simples dão suporte a sistemas cuja qualidade e abrangência da solução é muito inferior à humana, porém permitem a disseminação desse conhecimento a custos aceitáveis. Dessa forma, um número cada vez maior de aplicações de sistemas de conhecimento convive no dia a dia das empresas e dão fôlego a uma crescente indústria de software dedicada a este segmento.

Aplicações na WEB, por sua vez, respondem pela intensa pesquisa em modelagem conceitual e ontologias dos últimos anos. Ontologias são utilizadas para integrar vocabulários para comércio eletrônico, permitindo a construção de comparadores automáticos de serviços e preços (Kim, Choi *et al.*, 2005). Portais de conhecimento nos mais diversos domínios permitem a construção incremental de ontologias com o objetivo da definição de termos técnicos, processos e procedimentos em uma determinada área de aplicação. Por exemplo, o sítio do *Petrotechnical Open Standards Consortium*[5] busca definir um vocabulário padrão de referência para a troca de informação, construção de negócios, e armazenamento de dados na indústria de petróleo. Ontologias são aplicadas também para a integração de conteúdo e serviços disponíveis na WEB como a aplicação descrita neste artigo.

a. Projeto de Gestão de Conhecimento para monitoração de processo contínuo em indústria de papel

A produção de papéis filtrantes desenvolve-se através de um processo contínuo que exige permanente monitoração dos indicadores de manufatura, enquanto bobinas de papel são geradas à velocidade de 300 a 400 metros de papel por minuto. Durante a produção, são monitorados os parâmetros de composição da massa de celulose, umidade, calor, espessura, entre outras e realizadas ações corretivas para manter o processo dentro do padrão definido. No final da produção, o papel produzido é testado a intervalos regulares quanto à gramatura, umidade, porosidade, direção das fibras, alvura, espessura, presença de marcas no papel. Os

[5] *http://www.posc.org*

resultados desses parâmetros medidos em laboratório definem a classe do papel produzido, de acordo com uma tabela padrão da indústria.

As deficiências identificadas no estudo deste processo mostraram (Dummer 2003):

- Um pequeno número de técnicos tinham conhecimento do processo de monitoração expondo a empresa a problemas com o *turnover* de recursos humanos;
- As ações de correção do processo eram realizadas de maneira *ad hoc* pelos técnicos responsáveis, a empresa não tinha domínio sobre o processo de monitoração;
- Os padrões de classificação do papel produzido estavam desatualizados, a definição da classe do papel produzido era uma decisão subjetiva do laboratorista;
- A qualidade do papel produzido mostrava-se muito abaixo do desejável, gerando perdas de produto.

O projeto se desenvolveu em duas fases: a eliciação do conhecimento dos técnicos de monitoração e laboratoristas, e a análise dos resultados para geração de um conjunto de normas de ações de reparo e uma nova tabela de classificação de papel, com padrões definidos pela empresa.

A aquisição de conhecimento foi realizada com os quatro técnicos responsáveis pela monitoração do processo durante os quatro turnos, para evidenciar modificações das ações de reparo nos turnos diurnos e noturnos. Para cada um dos técnicos foram aplicadas as técnicas de observação sem intervenção e depois entrevistas concorrentes, solicitando que fossem justificada as ações de reparo. De posse dos resultados das entrevistas, foram elencadas as ações de reparo e aplicada à técnica de grade de repertórios, como mostrado na Tabela 1. A grade de repertórios (Kelly 1955) é uma técnica de aquisição de conhecimento que relaciona problemas e ações corretivas, solicitando ao agente que indique a eficiência destas ações. Os valores que preenchem a tabela indicam a prioridade de aplicação da ação para corrigir o indicador listado a cada linha (0 para baixa prioridade, 5 para mais alta

prioridade). A grade de repertório permitiu homogeneizar os procedimentos corretivos e garantir maior cobertura ao modelo de domínio, uma vez que nem todas as ações possíveis foram aplicadas durante as entrevistas realizadas. O resultado da aplicação da técnica foi uma tabela única definindo as ações de reparo necessárias a cada um dos cenários identificados no processo.

Tabela 1- Grade de Repertório para evidenciar fatores de aplicação de ações de reparo no processo de produção de papel da Classe A.

Ações corretivas / ***Indicador***	**Aumentar vapor**	**Reduzir velocidade**	**Aumentar pressão – 1ª. prensa**	**Aumentar pressão – 2ª. prensa**	**Baixar régua**
Gramatura	0	0	2	1	5
Espessura	1	0	3	2	5
Aspereza	2	1	1	0	4
Tração	4	5	1	1	2
Rasgo L	2	1	5	3	4
Rasgo T	0	2	2	0	1
Umidade C,M,T	5	3	4	4	0

Paralelamente, foram entrevistados os laboratoristas (chefe e um dos membros da equipe) responsáveis pela classificação do papel, somente nos turnos diurnos. Neste caso foram aplicadas as técnicas de observação e entrevistas concorrentes, seguidas de entrevistas estruturadas, onde os dados obtidos de entrevistas anteriores são utilizados para planejar métodos indiretos de extração de conhecimento. O objetivo era verificar a adequação dos padrões utilizados e de que forma eram ajustados pelos laboratoristas. O resultado do processo foi um estudo e redefinição dos padrões a serem utilizados na empresa.

O relatório do projeto de Gestão de Conhecimento descreve o mapeamento dos insumos de conhecimento utilizados no processo de produção e a nova definição dos padrões e controles de qualidade para a organização. Uma reengenharia das tarefas dos trabalhadores de conhecimento responsáveis pela qualidade do papel produzido (técnicos monitores do processo, laboratoristas de análise dos parâmetros de qualidade do papel produzido e classificadores dos tipos de papel) permitiu alinhar o controle do processo com as medidas de qualidade esperada. Como produto maior, a organização apropriou-se do processo de controle de qualidade diminuindo a fragilidade associada à dependência dos especialistas do processo.

8 CONCLUSÃO

Modelos de conhecimento demandam grande esforço no seu desenvolvimento, validação e manutenção. Esse esforço é plenamente recompensado quando o modelo gerado atende o suporte ao raciocínio e quando os custos do desenvolvimento podem ser divididos pela reutilização em outras aplicações. O objetivo da Engenharia de Conhecimento é oferecer o suporte metodológico para a construção de bases de conhecimento e ontologias que darão suporte a tarefas que demandam conhecimento nos mais diferentes domínios de problemas.

A abordagem de ontologias tem se mostrado útil para resolver muitos dos impasses na captura do conhecimento. Proporciona uma separação entre os problemas de modelagem conceitual e de representação sem deixar de fornecer uma solução integrada. Oferece uma opção viável de construção de bases de conhecimento ao oferecer a opção de reusabilidade do conhecimento em mais de uma aplicação, distribuindo desta forma os altos custos da construção de bases de conhecimento para propósitos específicos. Mais ainda, ontologias têm se mostrado uma solução para diferentes aplicações de integração de informações, construção de sistemas e projetos de Gestão de Conhecimento.

A modelagem abstrata do raciocínio de solução de problema progride a passos mais lentos do que a definição de modelos conceituais para capturar o conhecimento declarativo. A compreensão dos mecanismos do raciocínio e sua proposição de modelos potencialmente reusáveis ainda são os maiores empecilhos para a implementação de sistemas de solução de problemas eficazes. Métodos de solução de problemas mostraram-se um avanço importante na construção de raciocinadores automáticos, mas sua aplicação ainda não tem o alcance e a utilidade das ontologias.

A despeito de todo o sucesso da aplicação de ontologias nas mais diferentes aplicações e domínio, as fronteiras da pesquisa têm mostrado que a definição de uma linguagem ontológica robusta em termos semânticos e de processamento eficiente ainda está longe de ser alcançada. Novas necessidades, como o processamento na WEB e a transcendência dos idiomas têm aumentado as dificuldades em obter um padrão de representação.

Artigo recebido em 08 de outubro de 2012 e aceito para publicação em 16 de novembro de 2012

Abstract

The Knowledge Engineering field provides a series of methodologies and tools for knowledge modeling, with the aim of making knowledge more independent from people and allowing organizations and computer systems to profit from it. Thus, it fulfills the aims of Knowledge Management. Knowledge Engineering focuses on the creation of formal descriptions of problem-solving behavior, a process that is carried out by a rationalization of the observed solution. These models are defined in two levels: the knowledge level underlines the comprehension of the models by people; and the symbolic level aims at the mechanization of models, emphasizing computing efficiency. The knowledge, which is the object of the modeling task, is defined in Cognitive Science as a pattern of internal states that corresponds, one-by-one, to a state in the universe. The Computer Science, on the other hand, maintains that these patterns also can be internal states in a computer; implying that the inherent human abilities of keeping and using knowledge can be implemented in a computer. Still, the relation between the computer states and universe states can only be found by the human cognition. Knowledge is represented by a descriptive component, which represent the things in the world; and a dynamic component, which extracts more new information from the represented one. Nowadays, the most commonly used approached for representing the descriptive component are ontologies associated to domain models. Ontologies describe the meaning of the shared vocabulary in a domain, while knowledge models describe the instances and particularities of the considered problem. The dynamic component, in its turn, is described by the use of problem-solving methods. Ontologies have been showing the most successful approach in Knowledge Engineering; Semantic Web applications, expert systems, systems integration and knowledge portals can attest the usefulness of knowledge representation based on formal models, which are the aim of Knowledge Engineering.

KEY WORDS: Knowledge Engineering. Knowledge Management. Knowledge description. Ontology. Knowledge representation.

9 REFERÊNCIAS

Abel, M. **Estudo da perícia em petrografia sedimentar e sua importância para a engenharia de conhecimento.** (Tese de Doutorado). Programa de Pós-graduação em Computação, UFRGS, Porto Alegre, 2001.

Abel, M., L. S. Mastella, L. a. L. Silva, J. A. Campbell e L. F. D. Ros. **How to model visual knowledge: a study of expertise in oil-reservoir evaluation**. In: F. Galindo,

M. Takizawa*, et al* (Ed.). Database and Expert Systems Applications. Zaragoza , Spain: Springer-Verlag GmbH & Company KG, Berlin, Germany, v.3180, 2004. How to model visual knowledge: a study of expertise in oil-reservoir evaluation, p.455-464. (Lecture Notes in Computer Science)

Abel, M., L. a. L. Silva, L. F. De Ros, L. S. Mastella, J. A. Campbell e T. Novello. **PetroGrapher: Managing petrographic data and knowledge using an intelligent database application**. Expert Systems with Applications, v.26, n.1 SPECISS, p.9-18. 2004.

Anderson, B. F. **Knowing, learning and thinking**. In: (Ed.). Cognitive Psychology. London: Academic Press, 1975. Knowing, learning and thinking.

Audi, R.**_Epistemology: A Contemporary Introduction to the Theory of Knowledge**. London: Routledge, 1998.

Baader, F., D. Calvanese, D. Mcguinness, D. Nardi e P. Patel-Schneider. **The Description Logic Handbook: Theory, Implementation and Applications**: Cambridge University Press, 2004.

Barry, A. M. S. **Visual intelligence: perception, image and manipulation in visual communication**: State University of New York Press. 1997.

Bartalos, P., M. Barla, G. Frivolt, M. Tvarozek, A. Andrejko, M. Bielikova e P. Navrat. **Building an ontological base for experimental evaluation of Semantic Web applications.** 33rd Conference on Current Trends in Theory and Practice of Computer Science,SOFSEM 2007. Harrachov, Czech Republic: Springer Verlag. Jan 20-26 2007, 2007. p. 682-692 .

Bechhofer, S., F. V. Harmelen, J. Hendler, I. Horrocks, D. L. Mcguinness, P. F. Patel-Schneider e L. A. Stein. OWL Web Ontology Language Reference. W3C Recommendation 2004.

Beckett, D. **RDF/XML Syntax Specification** (Revised). Resource Description Framework (RDF). W3C Recommendation. 2004.

Berners-Lee, T., J. Hendler e O. Lassila. **The Semantic Web**. Scientific American, v.284, n.5, p.34-43. 2001.

Bodner, G. M. **Construtivism: a theory of knowledge**. Journal of Chemical Education, v.63, p.873-878. 1986.

Boff, L. H. **Conhecimento: fonte de riqueza das pessoas e das organizações.** Fascículo Profissionalização do Banco do Brasil. [S.l] 2000.

Boose, J. H. **Knowledge acquisition tools, methods, and mediating representations**. Japanese Knowledge Acquisition for Knowledge-based Systems Workshop, JKAW. Kyoto, Japan: IOS Press 1991.

Borst, W. N. **Construction of engineering ontologies**. (PhD). Centre for Telematics and Information Technology, University of Twente, Enschede, 1997.

Brachman, R. J. e H. J. Levesque. **Knowledge Representation and Reasoning.** San Francisco, USA: Morgan Kaufmann Publishers Inc., 2004.

Brachman, R. J. e J. G. Schmolze. **An overview of the KL-ONE Knowledge Representation System**. Cognitive Science, v.9, n.2, p.171-216. 1985.

Bray, T., J. Paoli, C. M. Sperberg-Mcqueen, E. Maler e F. Yergeau. Extensible **Markup Language (XML) 1.0**. W3C Recommendation, 2006.

Brickley, D. e R. V. Guha. **RDF Vocabulary Description Language 1.0: RDF Schema. Resource Description Framework (RDF)**. W3C Recommendation, 2004.

Buchanan, B. e E. Shortliffe. **Rule-based expert systems: the MYCIN experiments**. Reading: Addison-Wesley. 1984.

Burton, M., A., V. Bruce e P. J. B. Hancock. **From Pixel to People: A Model of Familiar Face Recognition**. Cognitive Science, v.23, n.1, p.1-31, 1999.

Castro, E. S. E., F. I. Victoreti, S. Fiorini, M. Abel e R. Price. **Um Caso de Integração de Gerenciamento Ágil de Projetos à Metodologia CommonKADS** . I Workshop de Gerenciamento de Projetos de Software. Florianópolis, 2008.

Chandrasekaram, B. **Generic tasks in knowledge-based reasoning: high level building blocks for expert system design**. IEEE Expert. Los Alamitos. 1**:** 23-30 p. 1986.

Clancey, W. J. **The knowledge level reinterpreted: modeling how systems interact**. Machine Learning. Boston**:** 285-291 p. 1989.

Drucker, P. F. **Sociedade Pós-Capitalista**. São Paulo: Pioneira. 1993.

Duda, R. O., P. E. Hart, N. J. Nilsson e G. L. Sutherland. **Semantic network representations in rule-based inference systems**. In: D. A. Waterman e F. Hayes-Roth (Ed.). Pattern Direct Inference Systems. New York: Academic Press, 1978. Semantic network representations in rule-based inference systems, p.203-221.

Farquhar, A., R. Fikes e J. Rice. **The Ontolingua Server: a tool for collaborative ontology construction**. International Journal of Human-Computer Studies . v.46, n.6, p.707-727, 1997.

Fensel, D. e E. Motta. **Structured development of problem solving methods**. IEEE Transactions on Knowledge and Data Engineering, v.13, n.6 Nov/Dec, p.913-932, 2001.

Fernández-Breis, J. T. e R. Martínez-Béjar. **A Cooperative Framework for Integrating Ontologies**. International Journal of Human-Computer Studies, 2002.

Fernandez-Lopez, M. e A. Gomez-Perez. **Overview and analysis of methodologies for building ontologies**. Knowledge Engineering Review, v.17, n.2, June 2002, p.129-156, 2002.

Flores, C. D., C. L. Hoher, M. Ladeira e R. M. Vicari. **Una Experiencia en el uso de redes probabilísticas para el diagnóstico médico : una experiencia brasileira** Informátic@Médica, v.2, n.8, March 2001, p.25-29, 2001.

Gennari, J. H., M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Crubezy, H. Eriksson, N. F. Noy e S. W. Tu. **The evolution of Protege: An environment for knowledge-based systems development**. International Journal of Human Computer Studies, v.58, n.1, p.89-123, 2003.

Gómez-Perez, A. **Survey on ontology development tools**. OGAI Journal (Oesterreichische Gesellschaft fuer Artificial Intelligence), v.22, n.2, p.6-16, 2003.

Gómez-Pérez, A. e V. R. Benjamins. **Overview of knowledge sharing and reuse components: Ontologies and problem-solving methods**. International Joint Conference on Artificial Intelligence(IJCAI-99), Workshop on Ontologies and Problem-Solving Methods (KRR5). V. R. Benjamins, B. Chandrasekaran*, et al*. Stockolm, Sweden, 1999.

Gómez-Pérez, A., M. Fernández-López e O. Corcho. **Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web**. London: Springer. 2007. (Advanced Information and Knowledge Processing)

Gruber, T. **Ontolingua: a mechanism to support portable ontologies**. Stanford: Knowledge System Laboratory, Stanford University, 1992.

Gruber, T. R. **A translation approach to portable ontology specifications**. Knowledge Acquisition, v.5, p.199-220, 1993.

Gruber, T. R. **Toward Principles for the Design of Ontologies Used for Knowledge Sharing**. International Journal of Human-Computer Studies, n.43, p.907-928, 1995.

Guarino, N. **Formal ontology, conceptual analysis and knowledge representation**. International Journal Human-Computer Studies, v.43, n.2/3, p.625-640, 1995.

Guarino, N. **Formal Ontology in Information Systems** Formal Ontology in Information Systems, FOIS'98. Trento, Italy**:** 6-8 June 1998.

Guarino, N. e C. Welty. **Evaluating ontological Decisions with Ontoclean**. Communications of the ACM v.45, n.2, February 2002, p.61-65, 2002.

Guarino, N. e C. A. Welty. **An overview of OntoClean. In: S. Staab e R. Studer** (Ed.). Handbook of Ontologies. Berlin: Springer 2004. An overview of OntoClean, p.151-171. (International Handbook on Information Systems)

Guizzardi, G. **Ontological Foundations for Structural Conceptual Models**. Enschede, The Netherlands: Universal Press, v.05-74, 2005. (CTIT PhD Thesis Series)

Hayes-Roth, F., D. A. Waterman e D. B. Lenat. **Building expert systems**. Reading, MA: Addison-Wesley,1983.

Hoffman, R. R. **The problem of extracting the knowledge of experts from the perspective of experimental psychology**. AI Magazine. 8**:** 53-67,1987.

Horrocks, I. **DAML+OIL: A Reasonable Web Ontology Language. Web Services, E-Business, and the Semantic Web,** : Springer Berlin / Heidelberg. 2002 (8th International Conference on Extending Database Technology)

K.Chaudhri, V., A. Farquhar, R. Fikes, P. D. Karp e J. P. Rice. **OKBC: a programmatic foundation for knowledge base interoperability**. . Fifteenth National/Tenth Conference On Artificial Intelligence/Innovative Applications Of Artificial Intelligence. Madison, USA: American Association for Artificial Intelligence, 1998.

Kalkan, V. D. **An overall view of knowledge management challenges for global business** Business Process Management Journal, v.14, n.3, p.390 - 400, 2008.

Kifer, M., G. Lausen e J. Wu. **Logical foundations of object-oriented and frame-based languages**. Journal of the ACM, v.42, n.4, p.741-843, 1995.

Kim, W., D. W. Choi e S. Park. **Product information meta-search framework for electronic commerce through ontology mapping**. Second European Sematic Web Conference, ESWC 2005 - The Semantic Web: Research and Applications. Crete, Greece Springer Verlag. 3532**:** 408-422, 2005.

Kosslyn, S. M. **Image and brain.The resolution of the imagery debate**. Cambridge: MIT Press, 1994

Lai, L. F. **A knowledge engineering approach to knowledge management.** Information Sciences, v.177, n.19, p.4072-4094, 2007.

Lenat, D. B. e R. V. Guha. **Building large knowledge-based systems. Representation and inference in the Cyc project**. Reading, Mass.: Addison-Wesley. 1990

Macgregor, R. M. **Inside The LOOM Description Classifier.** ACM SIGART Bulletin, v.2, n.3, p.88-92,1991.

Mastella, L. **Um modelo de conhecimento baseado em eventos para aquisição e representação de seqüências temporais em Petrografia Sedimentar**. (Dissertação de Mestrado). Programa de Pós-graduação em Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.

Mastella, L. S., M. Abel, L. C. Lamb e L. F. De Ros. **Cognitive Modelling of Event Ordering Reasoning in Imagistic Domains**. International Joint Conference in Artificial Intelligence. Edimburgh. 30 July to 5 August, 2005.

Mcdermott, J. **Preliminary steps towards a taxonomy of problem-solving methods**. In: S. Marcus (Ed.). Automating knowledge acquisition form expert systems. Netherlands: Kluwer Academic Publishers, 1988. Preliminary steps towards a taxonomy of problem-solving methods, p.225-255.

Minsky, M. **A framework to represent knowledge**. In: P. Winston (Ed.). The psychology of computer vision. New York: Mc-Graw-Hill, 1975. A framework to represent knowledge, p.211-277.

Newell, A. **The knowledge level**. Artificial Intelligence. Amsterdam. 18**:** 87-127, 1982.

Nonaka, I. e H. Takeuchi. **Criação de conhecimento na empresa**. São Paulo: Campus, 1997.

Piaget, J. **Biologia e Conhecimento: ensaio sobre as relações entre as regulações orgânicas e os processos cognoscitivos**. Petrópolis: Vozes, v.1. 1973. (Psicologia da Inteligência)

Polanyi, M. **The Tacit Dimension**. In: L. Prusak (Ed.). Knowledge in Organizations. Newton, MA: Butterworth-Heinemann, 1966. The Tacit Dimension, p.135-146

Preece, A., A. Flett, D. Sleeman, D. Curry, N. Meany e P. Perry. **Better knowledge management through knowledge engineering**. IEEE Intelligent Systems and Their Applications, v.16, n.1 , January/February, p.36-43, 2001.

Raggett, D., A. L. Hors e I. Jacobs. **HTML 4.01 Specification**. W3C. Recommendation, 1999.

Rautenberg, S., A. C. Gomes Filho, J. L. Todesco e F. Á. Ostuni-Gauthier. **Ferramenta ontoKEM: uma contribuição à Ciência da Informação para o desenvolvimento de ontologia**. Perspectivas em Ciência da Informação, v.15, n.1, Jan-Abr 2010, p.239-258, 2010.

Santin, C. E. **Construtos Ontológicos para Representação Simbólica de Conhecimento Visual**. (Dissertação de Mestrado). Programa de Pós-Graduação em Computação, UFRGS, Porto Alegre, 2007.

Schreiber, A. T. **The KADS approach to knowledge engineering**. Knowledge Acquisition. [S.l.]. 4**:** 1-4, 1992.

Schreiber, G., H. Akkermans, A. Anjewierden, R. D. Hoog, N. Shadbolt, W. V. D. Velde e B. Wielinga. **Knowledge engineering and management:The CommonKADS Methodology**,Cambridge: The MIT Press. 2000.

Silva, L. a. L. **Aplicando métodos de solução de problemas em tarefas de interpretação de rochas**. (Dissertação de Mestrado). Programa de PG em Ciência da Computação, UFRGS, Porto Alegre, 2001.

Steels, L. **Components of expertise**. AI Magazine. [S.l.]. 11**:** 28-49, 1990.

Stokes, M. **Managing Engineering Knowledge. MOKA: Methodology for Knowledge-Based Engineering Applications**: ASME Press, v.1. 2001.

Studer, R., V. R. Benjamins e D. Fensel. **Knowledge Engineering: Principles and Methods**. IEEE Transactions on Data and Knowledge Engineering, v.25 (1-2), p.161-197, 1998.

Stutt, A. e E. Motta. **VITAL - A methodology-based workbench for KBS life cycle suppor**t. ESPRIT II. [S.l.]. 1994. (P5365)

Ullmann, S. **Semantics: An Introduction to the Science of Meaning**. Oxford: Rowman & Littlefield, 1979.

Velde, W. V. **Issues in knowledge level modeling**. In: J. M. David, J. P. Krivine*, et al* (Ed.). Second generation expert systems. Berlin: Springer Verlag, 1993.

Wielinga, B. J. e A. T. Schreiber. **Conceptual modeling on large reusable knowledge bases**. In: K. Von Luck e H. Marburger (Ed.). Management and processing of complex data structures. Berlin: Springer -Verlag, 1994. Conceptual modeling on large reusable knowledge bases, p.181-200. (Lecture Notes in Computer Science v.777)

View publication stats