

OntoSIGF - um sistema de informação geográfica baseado em ontologias aplicado à área florestal

OntoSIGF - a geographic information system based on ontologies applied to forestry

Bayron Thiengo Quinelato¹

Dornélio Mori Junior²

Rober Marcone Rosi³

Resumo

O desenvolvimento dos sistemas de navegação e posicionamento global, tecnologias *web*, *software*, e *hardware* possibilitaram a rápida expansão do uso de Sistemas de Informação Geográfica (SIG). A informação é sempre distribuída e principalmente heterogênea. Contudo, especialistas de domínio usam conceitos e terminologias específicas e usam diferentes parâmetros para descrever o modelo de um conceito. Esta heterogeneidade semântica das fontes de dados causa sérios problemas, que podem ser resolvidos aplicando o conceito de ontologias. O propósito deste artigo é o de descrever uma arquitetura baseada em ontologia capaz de registrar, descrever, organizar e integrar os conhecimentos de um domínio geográfico florestal. Este modelo ontológico descreve os conceitos das informações geográficas e suas relações, facilitando o acesso às informações em arquivos e bases de dados, independente de onde essas informações estejam armazenadas.

Palavras-chave: OntoSIGF; ontologia; metadados; sistemas de informação geográfica.

Abstract

The development of global positioning and navigation systems, web technology, software and hardware has fostered rapid expansion in the use

1 FAESA - Rua Anselmo Serrat, 199 - Ilha de Monte Belo, Vitória, ES. CEP: 29053-250; e-mail: bayrontq@gmail.com

2 FAESA - Rua Anselmo Serrat, 199 - Ilha de Monte Belo, Vitória, ES. CEP: 29053-250; e-mail: dmj.tanaka@gmail.com

3 FAESA - Rua Anselmo Serrat, 199 - Ilha de Monte Belo, Vitória, ES. CEP: 29053-250; e-mail: rober@faesa.br

of Geographic Information Systems (GIS). The available information is always distributed and mostly heterogeneous. However, domain experts use the concepts and terminology specific and use different parameters to express their model of a concept. Such semantic heterogeneity of data sources causes serious problems, which may be resolved by ontologies. The purpose of this article is to describe an ontology-based architecture to register, describe, organize and integrate the geographical forest knowledge. This ontological model describes the geographic information concepts and their relationships, facilitating the information access in archives and databases, regardless of where such information is stored.

Key words: OntoSIGF; ontology; metadata; geographic information system.

Introdução

O gerenciamento da informação é a chave para planejamentos e tomada de decisões eficazes em qualquer domínio. No caso da informação geográfica esse gerenciamento é de uma complexidade maior do que na maioria das outras informações. Com a evolução dos equipamentos de obtenção de informações geográficas e o dinamismo na troca de informações geradas pelas modernas redes de computadores, o volume e distinção dos formatos das informações cresceram muito mais do que a interoperabilidade entre os SIGs. Um dos problemas mais graves enfrentados pelos sistemas de informação é a falta de padronização dos dados, o que torna muito difícil o intercâmbio de dados e informações entre sistemas distintos, gerando duplicação de esforços e disponibilizando informações de forma menos abrangente, menos acessível e com custos elevados.

A formalização do conhecimento em uma linguagem computacional é um

desafio para a área de computação. Neste aspecto, a utilização de ontologias na descrição semântica de um determinado vocabulário vem se destacando, pois proporciona um entendimento amplo das características e propriedades das classes pertencentes a um domínio, assim como seus relacionamentos. Ontologias vêm sendo desenvolvidas para estabelecer consenso sobre o significado de conceitos específicos de diversos domínios do conhecimento e assim resolver problemas de interoperabilidade. Devido a estes fatores, a ontologia tem sido bastante aplicada no domínio geográfico.

Este artigo apresenta uma proposta de arquitetura para um sistema de apoio a um domínio geográfico florestal. Tendo em vista a diversidade de dados geográficos gerados, foi criada uma ontologia de domínio, a qual registra todo o conhecimento do domínio geográfico de empresas que trabalham com manejo de áreas florestais. Dessa forma, ao possibilitar o registro do conhecimento, facilita o entendimento do mesmo e evita

que este conhecimento se perca, além de padronizar os conceitos para a empresa em questão e conseqüentemente, para toda a informação gerada a partir dos sistemas baseados na ontologia proposta.

Também foi desenvolvido um protótipo baseado na ontologia proposta que efetua a busca por mapas comparando os conceitos digitados pelo usuário com as palavras-chave contidas nos metadados dos mapas e com os conceitos registrados na ontologia, utilizando-se as linguagens *Java* e *Web Ontology Language (OWL)* e as ferramentas de apoio pesquisadas: Protégé, Pellet, Jena, CMap e Jude, que combinadas foram capazes de contornar o problema da falta de um modelo conciso de operacionalização de ontologias.

Ontologias

Ontologia é uma antiga disciplina que vem desde o estudo feito por Aristóteles sobre as categorias e a metafísica. O primeiro uso de ontologias em sistemas de computação foi através dos sistemas de inteligência artificial. Atualmente, diversas áreas usam ontologias. Na área de Filosofia o conceito de ontologia se refere a um sistema particular de categorias de acordo com certa visão do mundo, não tendo uma linguagem específica para representação. Na área de Inteligência Artificial e, em geral, para toda a comunidade da Ciência da Computação, o conceito de ontologia se refere a artefatos de engenharia, constituídos por um vocabulário específico usado para descrever certa realidade. Gruber (1992) propôs que “uma ontologia é uma especificação explícita de uma conceitualização”. Guarino (1998) estende essa definição dizendo que uma onto-

logia é na verdade uma especificação parcial e explícita que tenta, da melhor forma possível, aproximar a estrutura de mundo definida por uma conceitualização. Uma ontologia, portanto, passa a ter compromisso apenas com a consistência em um determinado domínio e não com a completude.

A ontologia implementada no OntoSIGF visa descrever, organizar e integrar as diferentes bases de dados geográficas possibilitando a interoperabilidade destas bases, facilitando o reuso do conhecimento, minimizando o retrabalho e, conseqüentemente, reduzindo os custos de desenvolvimento de uma solução baseada em SIG. Além de possuir uma interface que permita sua ampliação e manutenção, registrar formalmente o conhecimento produzido pelas entidades ou empresas de manejo de áreas florestais e permitir a integração do conhecimento em um ambiente de desenvolvimento de sistemas de informação geográfica florestal.

O OntoSIGF

No domínio de aplicação estudado neste trabalho, apesar da grande quantidade de dados geográficos disponíveis, não existe um formato padrão (planilhas, arquivos, banco de dados etc.) de armazenamento, dificultando a integração e utilização. Além disso, dados geográficos possuem características que dificultam o seu gerenciamento, tais como: grande volume de dados a ser processado; heterogeneidade de tipos de dados e de ambientes (*software* e *hardware*); estrutura complexa, com características espaço-temporal incertos e altamente

distribuídos (cada informação ou conjunto de dados pode estar localizado em locais diferentes, dificultando ainda mais sua recuperação e integração). Somam-se a isto os principais problemas enfrentados pelos usuários na busca de informações geográfica: os dados não existem ou são insuficientes; a qualidade dos dados recuperados é difícil de ser medida.

As bases de dados geográficas são atualizadas com uma periodicidade muito grande, mapas são gerados para os usuários, para análise de informações, pesquisa, etc, produzindo diariamente uma quantidade muito grande de dados geográficos. Todos estes dados são gerados pelos técnicos e analistas cartográficos que buscam essas informações em três bases de dados distintas: estrutura de pastas e arquivos, base de dados geográficos e mídias externas como CDs, DVDs, entre outros.

Os dados são gerados de acordo com as demandas, sendo uma dessas demandas a solicitação gerada pelos usuários da cartografia que são as áreas de fomento, terras, silvicultura, etc. Depois de gerada, a solicitação é enviada para a cartografia que realiza o atendimento. O atendimento se dá através da leitura e entendimento da solicitação e captura dos dados. Estes dados são capturados em uma estrutura de pastas, arquivos e banco de dados geográfico. Como esta estrutura possui um tipo de padronização ou nomenclatura informal e setorizada, os dados são armazenados conforme o entendimento do usuário da cartografia, ampliando a dificuldade de recuperação posterior. Os técnicos e analistas cartográficos muitas vezes não sabem se este dado já foi produzido ou não, e como não existe um mecanismo adequado

de busca e pesquisa para encontrar o dado, ele é pesquisado manualmente.

Uma vez que o dado solicitado não seja encontrado, um novo dado é produzido, surgindo problemas como duplicação de dados nas bases, falta de controle no dado gerado e demora no atendimento das solicitações. Tendo em vista este cenário, o desafio está em como organizar os dados geográficos de modo a permitir que esta base de dados evolua e que existam mecanismos adequados de busca e pesquisa destes dados.

O OntoSIGF foi desenvolvido para servir como base para a solucionar os problemas de registro do conhecimento descritos acima, facilitar o entendimento do mesmo e evitar que este conhecimento se perca. Também busca padronizar os conceitos para a empresa em questão e conseqüentemente, para toda a informação gerada a partir dos sistemas baseados na ontologia proposta para o domínio florestal. Objetivando validar a ontologia e exemplificar o desenvolvimento de um sistema baseado em ontologia, foi desenvolvido um protótipo composto por um conjunto de ferramentas de apoio que possibilitaram colocar a ontologia em produção. Constitui este conjunto de ferramentas o Protégé, Pellet, Jena, CMap e Jude.

O Protégé foi implementado em *Java* e é um ambiente integrado, independente de plataforma e que provê um conjunto de funcionalidades para a construção de Ontologias e bases de conhecimentos. O Protégé oferece uma grande variedade de ações e estruturas que facilitam a construção, visualização e edição de Ontologias em vários formatos, tais como: RDF(S), OWL e XML Schema.

Pellet é um *reasoner* para OWL DL, desenvolvido em *Java* pela University of Maryland's Mindswap Lab e posteriormente tornou-se um projeto open source. O Pellet utiliza o protocolo conhecido como DIG que permite a interação de *reasoners* com o Protégé. A única configuração necessária é a definição de uma porta para comunicação em ambos os aplicativos. O núcleo do Pellet consiste em um quadro de algoritmos que inferem e verificam a consistência da Ontologia.

Jena é um *framework, open source*, escrito em *Java* e desenvolvido pela Hewlett-Packard (HP). O Jena disponibiliza uma API que permite uma aplicação manipular ontologias. Seu objetivo é proporcionar um *framework* na linguagem *Java* que dê suporte à utilização da *Web* semântica por qualquer aplicativo capaz de utilizá-lo. Este suporte inclui recursos para manipulação de RDF, RDFS, OWL e DAML+OIL.

O IHMC CMap foi desenvolvido pelo *Institute for Human and Machine Cognition* (IHMC). É um *software* de criação de mapas conceituais, estes possuem componentes comuns com a ontologia, por exemplo, captura de conhecimento, o compartilhamento e representação da informação.

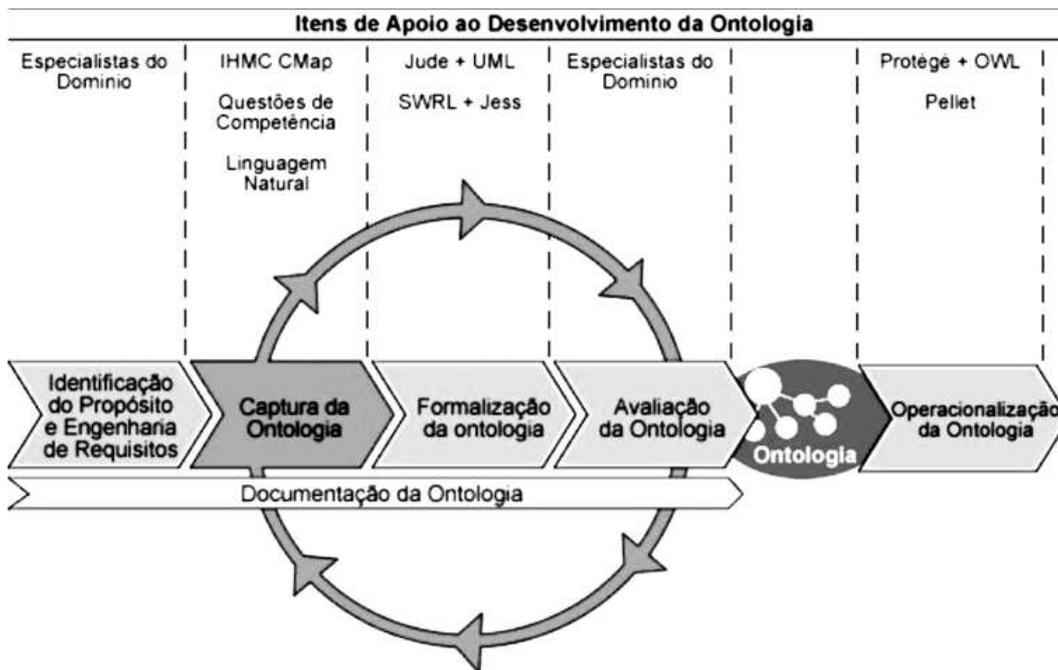
Para atingir os objetivos pretendidos na elaboração da ontologia, o projeto desenvolveu uma metodologia baseada principalmente na metodologia de construção de ontologias *Systematic Approach for Building Ontologies* (SABiO), proposta por Falbo, e na metodologia *On-To-Knowledge*, proposta por Staab. A figura 1 ilustra as fases de desenvolvimento da ontologia e os itens de apoio a cada uma destas fases.

A definição do domínio da ontologia tem por objetivo estabelecer o escopo dos conceitos geográficos que a ontologia irá atender. Para isto utilizam-se questões de competência relacionadas ao domínio e que devem ser respondidas pelos especialistas do domínio, além da consulta de documentos, dicionários geográficos e bases de dados.

A análise dos requisitos é a etapa que detalha este escopo inicial e consiste de várias reuniões com especialistas do setor cartográfico, e na análise de documentos, dicionários geográficos e bases de dados da empresa.

Para possibilitar uma visão mais ampla dos conceitos e suas relações e para facilitar a comunicação com os especialistas do domínio, foi elaborado um mapa conceitual com a ferramenta CMap. A partir das informações obtidas com as questões de competência, foi gerada uma versão inicial do mapa conceitual. A ontologia passou por diversas análises dos especialistas, até que se chegou a uma versão totalmente validada pelos especialistas de domínio consultados. A metodologia apresentada propõe o uso de uma linguagem gráfica para expressar ontologias e de uma linguagem formal, para formalizar seus axiomas. A linguagem gráfica adotada foi a *Ontology UML Profile* (OUP) que é um perfil *Unified Modeling Language* (UML) mais simplificado, no qual constam apenas os elementos que representam conceitos utilizados no desenvolvimento de ontologias como classes, indivíduos e propriedades. A ferramenta de modelagem usada para criação do modelo visual da ontologia proposta foi o JUDE, por ser uma ferramenta bastante difundida e

Figura 1. Metodologia de desenvolvimento de ontologias



com total suporte à modelagem UML. A formalização dos axiomas é feita a partir da captura dos axiomas, os quais são responsáveis por modelar as sentenças verdadeiras da ontologia e especificar as restrições sobre a sua interpretação. Sua captura também é feita através de consultas aos especialistas do domínio e a documentos e base de dados da cartografia.

Validados os conceitos e seus axiomas, faz-se a formalização da ontologia. Para isto utilizou-se o Protégé, que permite a criação e edição de ontologias. Depois de formalizada, a ontologia é exportada para a linguagem OWL, que é uma linguagem semântica de representação de ontologias recomendada pelo consórcio *World Wide Web Consortium* (W3C).

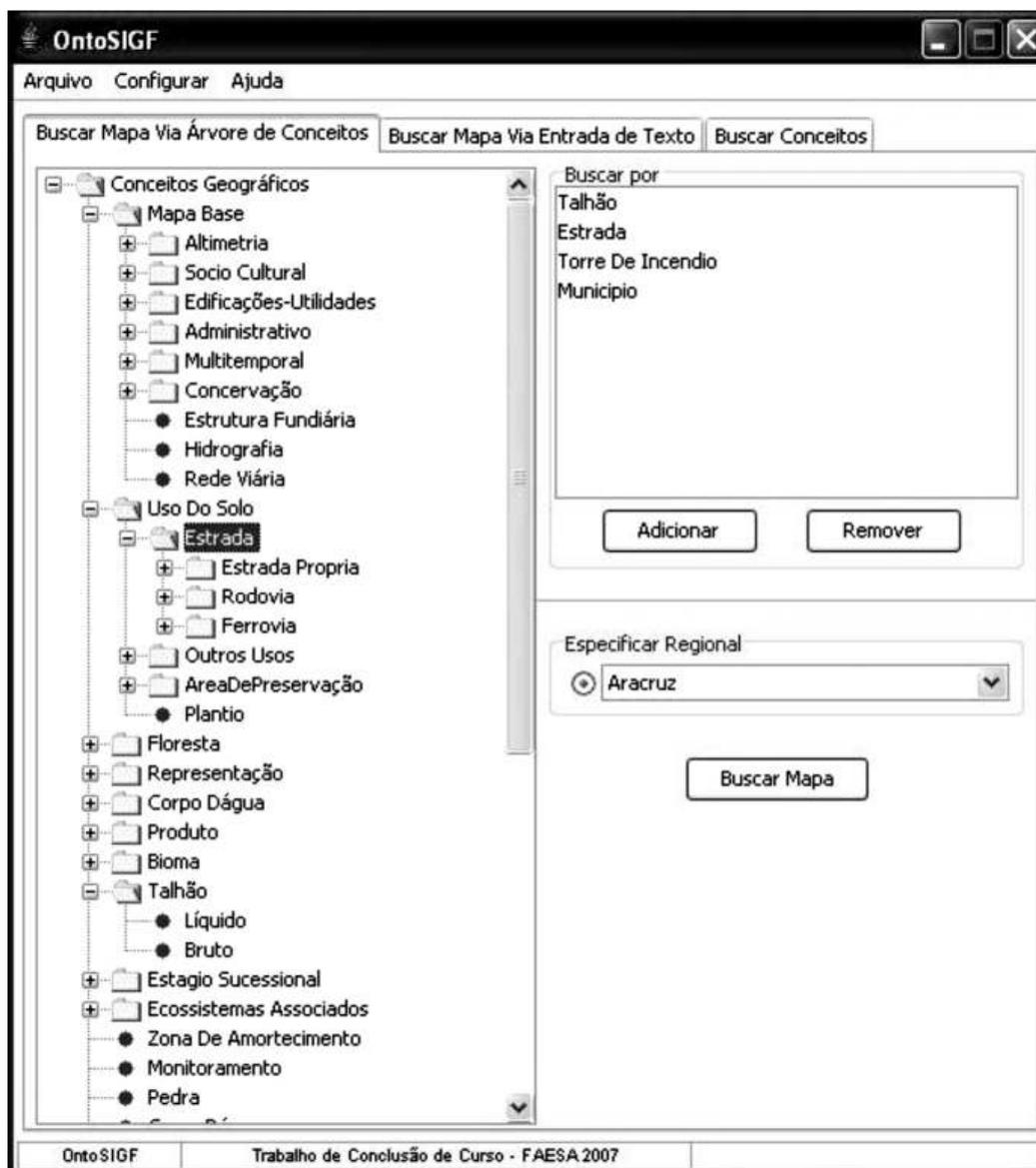
Para a implementação do protótipo baseado na ontologia desenvolvida, foi

necessária a utilização de um *framework* para o desenvolvimento de aplicações que manipule ontologias. No OntoSIGF foi utilizado o *framework* Jena. O protótipo foi implementado na linguagem *Java*, que também é a linguagem na qual o Jena foi desenvolvido.

O protótipo implementado consiste em um sistema baseado na ontologia desenvolvida e que possui uma funcionalidade básica de busca por mapas através da comparação dos conceitos contidos na ontologia com os conceitos digitados ou selecionados pelo usuário e com os conceitos contidos nos metadados dos mapas. A figura 2 demonstra a arquitetura proposta para o sistema OntoSIGF.

Pode-se notar pela figura 3, que a captura dos conceitos e sua formalização a partir da implementação da ontologia pôde gerar melhorias nos

Figura 3. Tela do protótipo demonstrando árvore de consulta por conceito ontológico



a descrição dos dados e promove a interoperabilidade no nível semântico.

Venâncio utiliza um referencial ontológico comum para explicitar relações de agregação e especialização entre entidades geográficas em um sistema de navegação em mapas dirigida por ontologia (OntoCarta). Isso contribui

para a integração de dados, pois permite descrever e identificar cada objeto geográfico de acordo com uma conceitualização consensual.

Chaves propôs o modelo *Geographic Knowledge Base* (GKB) para representar a ontologia física de Portugal. O GKB é um meta-modelo capaz de

integrar o conhecimento geográfico coletado de diversas fontes.

Segundo Vckovski, no caso dos SIGs, as ontologias são utilizadas para solucionar problemas de representação espacial de objetos geográficos e para promover a interoperabilidade entre SIGs com diferentes arquiteturas e formatos de dados. Porém, a interconexão entre estes sistemas de informações geográficas está longe de ser alcançada.

Conclusões e Perspectivas Futuras

Este trabalho descreve a construção de uma ontologia que registra e organiza o conhecimento do domínio geográfico de uma empresa que trabalha com manejo de áreas florestais. Esta ontologia permite registrar o conhecimento deste domínio, descrevendo seus conceitos, relações e demais características.

Além da ontologia, foi apresentada uma metodologia para construção desta ontologia, bem como uma arquitetura para construção do sistema baseado na ontologia do domínio proposto.

Como a área de aplicações baseadas em ontologias é muito nova, espera-se que este trabalho possa contribuir com a construção de novos projetos, visto que ele apresenta uma proposta completa de desenvolvimento de sistemas baseados em ontologias, desde sua concepção

até a fase de codificação e implantação. O OntoSIGF propõe um processo de busca e geração de mapas da cartografia baseado na ontologia proposta, além de contribuir para implementações futuras de sistemas de informação geográfica florestal através de sua ontologia.

As perspectivas futuras para este trabalho envolvem o refinamento da ontologia desenvolvida, tornando-a o mais genérica possível, com a finalidade de ser compatível ao máximo com o domínio geográfico de diferentes em-presas que trabalham com manejo de áreas florestais e o desenvolvimento de ontologias de domínio que possam interligar setores específicos da em-presa que fazem interface direta ou indi-retamente com a cartografia, para que o SIG possa compreender e responder às solicitações feitas por esses clientes ao setor cartográfico. Também envolve o desenvolvimento de um SIG Florestal voltado para a *web*, mais complexo e que segue as características básicas do protótipo proposto para atender a demanda de geração de mapas de grandes em-presas que trabalham com manejo de áreas florestais.

Agradecimentos

Ao apoio estrutural e intelectual oferecido pela FAESA, Aracruz Celulose e INFLOR.

Referências e bibliografia de apoio

GUARINO, N. *Formal Ontology and Information Systems*. In: FOIS 98, 1998, v.6, n.8, p.3–15, Trento, Itália. . .

GRUBER, T. A translation approach to portable ontology specifications. *Technical Report KSL 92-71*. Knowledge Systems Laboratory - Stanford University, Stanford, CA, 1992.

FALBO, R.A. *Integração de conhecimento em um ambiente de desenvolvimento de software*. Rio de Janeiro, 1998. Tese (Doutorado em Engenharia de Sistemas de Computação). COPPE/ Universidade Federal do Rio de Janeiro.

BROCKMANS, S.; HAASE, P.; HITZLER, P. *A Metamodel and UML Profile for Rule-extended OWL DL ontologies*. *The Semantic Web: Research and Applications*, Springer, Budva, v. 4011 of LNCS, p. 303-316, 2006.

GUIZZARDI, G. *Ontological Foundations for Structural Conceptual Models*, Twente (Holanda), 2005. Tese (Doutorado em Ph.D. in Computer Science) Twente University of Technology, Holanda.

FONSECA, F. et al. *Using Ontologies for Integrated Geographic Information Systems*. INPE - Instituto Nacional de Pesquisas Espaciais, São Paulo, Brasil, 2002.

LIMA, P; CÂMARA, G; QUEIROZ, G. *GeoBR: Intercâmbio sintático e semântico de dados espaciais*. INPE - Instituto Nacional de Pesquisas Espaciais, São Paulo, Brasil, 2002.

VENANCIO, L. R; FILETO, R; MEDEIROS, C. B. *Aplicando Ontologias de Objetos Geográficos para Facilitar Navegação em GIS*. In: SIMPÓSIO BRASILEIRO DE GEOINFORMÁTICA (GEOINFO), 5º, São Paulo, Brasil, 2003. Anais...

CHAVES, M. S; SILVA, M. J; MARTINS, B. A *Geographic Knowledge Base for Semantic Web Applications*. In: SIMPÓSIO BRASILEIRO DE BANCO DE DADOS (SBBDD), 20º, Mina Gerais, Brasil, 2005. Anais...

VCKOVSKI, A. *Special Issue: Interoperability in GIS*. *International Journal of Geographical Information Science*, Londres, Inglaterra, v. 12, n. 4, p. 297-298, 1998.

STAAB, S.; SCHNURR, H.P.; STUDER, R.; SURE, York. *Knowledge Processes and Ontologies*. 2001. [online]. <[www.aifb.uni-](http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2001_knowledge_processes_and_ontologies.pdf)

[karlsruhe.de/WBS/ysu/publications/2001_knowledge_processes_and_ontologies.pdf](http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2001_knowledge_processes_and_ontologies.pdf)>. Acesso em 30 de nov., 2007.

W3C Recommendation – Web Ontology Language. Disponível em: <<http://www.w3.org/TR/owl-features/>> Acesso em 14 de jul., 2008.