

Uma introdução à *Web Semântica* no domínio dos Sistemas de Informações Geográficas

Angelo Augusto Frozza, Rodrigo Gonçalves

{frozza,rodrigog}@inf.ufsc.br
Universidade Federal de Santa Catarina – UFSC
Florianópolis – Santa Catarina

Resumo. A *Web* apresenta-se atualmente como uma das fontes de pesquisa mais utilizadas. Porém possui algumas limitações que impedem a realização de buscas mais precisas pelas informações desejadas. Visando corrigir este quadro, introduziu-se o conceito de *Web Semântica*. Porém, boa parte da *Web* ainda não incorporou os recursos necessários para que a *Web Semântica* se concretize. Similarmente, os sistemas de Informações Geográficas (SIG) na *Web* (*WebGIS*) apresentam a mesma limitação, acrescida das particularidades dos dados geográficos. Em razão das características similares com outras aplicações na *Web* (possuem uma grande gama de dados porém pouca significância para os mesmos), dos *WebGIS* podem beneficiar-se dos avanços da *Web Semântica*. É sobre este novo quadro dos *WebGIS* que trata este artigo.

1. Introdução

Chamam-se Sistemas de Informações Geográficas (SIG) os sistemas que realizam o tratamento computacional de dados geográficos. A principal diferença de um SIG para um sistema de informação convencional é sua capacidade de armazenar tanto os atributos descritivos como as geometrias dos diferentes tipos de dados geográficos. Entrada bibliográfica.

As principais características de SIGs são Entrada bibliográfica:

- Inserir e integrar, numa única base de dados, informações espaciais provenientes de meio físico-biótico, de dados censitários, de cadastros urbanos e rurais e outras fontes de dados como imagens de satélite e GPS;
- Oferecer mecanismos para combinar as várias informações através de algoritmos de manipulação e análise, bem como para consultar, recuperar e visualizar o conteúdo da base de dados geográficos.

Atualmente, uma das limitações para a maior popularização dos SIGs corresponde às diferenças existentes entre os formatos de dados usados, geralmente proprietários, o que limita a interoperabilidade entre SIGs distintos. Com o advento da Internet, que permite a disseminação de dados de forma ampla e para um público heterogêneo, as informações existentes nos SIGs também foram levadas para a *Web*. A partir de então, percebeu-se a oportunidade de usar as tecnologias da *Web* para a integração de SIGs diferentes e, assim, ampliar as possibilidades de consulta até então disponíveis, originando os sistemas *WebGIS*.

A partir da necessidade de compartilhar informação de forma eficiente com um público com interesses diversos, uma das soluções mais aceitas é o emprego de ontologias

em sistemas de informação. Um sistema de informação pode ser concebido como um mecanismo de comunicação entre duas partes: o produtor e o usuário. Para que funcione, é necessário que haja uma concordância entre os conceitos das partes. Numa perspectiva mais geral, seu sucesso depende da existência de uma comunidade que compartilhe as definições utilizadas para construí-lo. Entrada bibliográfica. Estas definições são representadas por meio de uma *ontologia de aplicação*.

No caso de dados geográficos, uma ontologia de aplicação, ou geo-ontologia, tem dois tipos básicos de conceitos Entrada bibliográfica:

- a) conceitos que correspondem a fenômenos físicos do mundo real, ou conceitos físicos;
- b) conceitos criados para representar entidades sociais e institucionais, ou conceitos sociais.

Porém, na maior parte dos sistemas de informação atuais, as ontologias de aplicação não estão explicitadas, o que reduz o potencial de compartilhamento da informação. A explicitação das ontologias de aplicação tem por base as propostas recentes da *Web Semântica* Entrada bibliográfica e de propostas de padrões como OWL. Para tanto, existem vários sistemas disponíveis na Internet para criação e gestão de ontologias, como o Protegé Entrada bibliográfica. Para dados geográficos, o consórcio OGC (*Open Geospatial Consortium*) propôs o formato GML Entrada bibliográfica como mecanismo de transporte de dados geográficos.

O desenvolvimento de *WebGIS* é um tema de pesquisa bastante atual, sem previsão para consolidação deste tipo de tecnologia. Nesta linha, pode-se citar dois tópicos de pesquisa em aberto Entrada bibliográfica:

- a) a necessidade de uma forma canônica razoável que proponha requisições de dados geoespaciais, o que leva aos estudos sobre interoperabilidade entre SIGs;
- b) a necessidade de métodos para avaliar a semântica de fontes de dados disponíveis, a fim de ver se sua estrutura semântica pode ser explorada para qualquer consulta geoespacial. Esta linha de pesquisa deu origem à *Web Semântica Geoespacial*.

Na prática, a grande vantagem de se unir Sistemas de Informações Geográficas (SIG) e a *Web Semântica* é a possibilidade de se poder indicar com precisão a localização geográfica de um local. Por exemplo, um usuário fazendo uma busca por informações sobre São Paulo (*cidade*), não está interessado em *sites* sobre o *Estado* de São Paulo. Apesar de ser um exemplo bastante simples, é de grande complexidade e dimensão, uma vez que há muitos povos, cidades e lugares que têm o mesmo nome.

O ponto de partida para a recuperação da informação na *Web Semântica Geoespacial* é uma *requisição geoespacial* que tenha claramente definidas as semânticas associadas com cada elemento. Para tanto, é necessário o desenvolvimento e disponibilização de ontologias espaciais e terminológicas, representadas de tal forma que possam ser usadas tanto para processamento por máquinas como para o entendimento por pessoas. A partir disso, processam-se consultas geoespaciais contra estas ontologias. A avaliação dos resultados baseia-se na comparação entre as semânticas da informação necessária e as semânticas disponíveis das fontes de informação e sistemas de busca Entrada bibliográfica.

2. WebGIS e interoperabilidade

A interoperabilidade entre Sistemas de Informações Geográficas visa possibilitar o trabalho conjunto dos mesmos de forma a permitir consultas sobre as informações neles contidas. Um sistema SIG isolado por si só já permite uma série de consultas e a recuperação de uma grande quantidade de informações. Porém, se integrado a outros SIGs, pode melhorar a qualidade e riqueza das informações que retorna, além de aumentar as possibilidades de consultas realizáveis (através do cruzamento de dados, por exemplo). A interoperabilidade busca dotar as aplicações de recursos para interpretar e processar dados oriundos de diversas fontes, bem como localizar e acessar as próprias fontes

Entrada bibliográfica.

A integração dos SIG pode ser considerada em três níveis Entrada bibliográfica:

- *Sintático* - Compatibilização de tipos de dados (*strings*, inteiros etc.);
- *Estrutural* - Converter a estrutura dos dados de um sistema para uma estrutura compatível com outro sistema;
- *Semântica* - Compatibilização entre representações segundo paradigmas diferentes (O.O. e relacional) - heterogeneidade semântica genérica - ou compatibilização de conceitos e contextos de dados - heterogeneidade semântica contextual.

Os sistemas de SIG atuais possuem diversas formas de organizar e codificar os dados (geralmente em formatos de dados proprietários), dificultando o processamento de consultas sobre diversos sistemas de SIG simultaneamente, para posterior cruzamento de informações e análise. Porém, um esforço do *Open GIS Consortium* definiu um padrão de transporte de dados, a linguagem *Geography Markup Language* - GML (baseada no XML), que define a representação de aspectos geográficos dos SIGs. Porém, a GML não garante a correta interpretação dos dados, pois não considera semânticas associadas aos objetos geográficos representados pela mesma Entrada bibliográfica. Logo, um ponto que permanece em aberto refere-se à organização dos dados dos sistemas SIG não relacionados diretamente a mapas (dados geográficos), mas a características destes dados geográficos (atributos, estatísticas etc.). Não existe ainda um padrão de representação ou uma forma de consulta uniforme.

Estes problemas de integração entre os SIG são semelhantes aos existentes quando se busca a integração de fontes de dados na *Web*. Logo, as soluções para esta última tendem a aplicar-se para os SIG também, melhorando os recursos dos *WebGIS* existentes. Por outro lado, realizar intercâmbio de dados em SIGs não é uma tarefa simples, devido a complexidade da informação geográfica envolvida, ocorrendo incompatibilidades em vários níveis Entrada bibliográfica.

3. WebGIS e Web Semântica

O grande fator limitante no acesso aos dados geográficos são as diferenças de entendimento entre comunidades de usuários distintas. Diferentes visões da realidade geográfica sempre existirão em pessoas com culturas diferentes. Assim, pode-se conviver com estas diferentes formas de conhecimento sobre a realidade pela criação de mecanismos que implementem e combinem diferentes visões, buscando a interoperabilidade pela equivalência semântica dos conceitos entre sistemas distintos

Entrada bibliográfica.

Os sistemas de *WebGIS* podem ou não fazer uso de recursos que permitam uma utilização semântica dos mesmos. No caso de sistemas de pesquisa em mapas geográficos, podem-se utilizar ontologias que permitam distinguir conceitos homônimos na hora de localizá-los no mapa (*Cidade* de São Paulo em vez de *Estado*). Outros SIGs podem usar aspectos semânticos, como ontologias, para organizar e permitir a criação e edição de mapas e conjuntos de informações geográficas.

Os *WebGIS* podem prover o acesso e manipulação das informações contidas nos SIGs através de interfaces *Web* acessíveis por entidades automatizadas (agentes) visando consultas e cruzamento de informações entre bases de dados geográficas distintas. Neste tipo de *WebGIS* a *Web Semântica* tem sua maior influência através do uso de tecnologias como RDF e ontologias para representar e apresentar as informações.

A maioria dos *WebGIS* existentes permite apenas a construção e publicação de mapas digitais na *Web*. Funções mais avançadas como modelagem, atualizações e análise de fontes de dados heterogêneas não estão implementadas. Em parte, isto deve-se ao fato destes serviços fundamentarem-se na primeira geração da *Web*, baseada em HTML. Entrada bibliográfica.

Porém, com a introdução da *Web Semântica*, conjugada ao uso de XML, RDF, OWL e ontologias, torna-se possível a implementação das funções descritas acima. Um aspecto se observar é como fazer a integração semântica dos dados? Uma ontologia de domínio pode ser utilizada para descrever os conceitos envolvidos nos *WebGIS* como a chave para a resolução deste problema. Entrada bibliográfica.

O XML é um dos principais métodos para troca de informações dos *WebGIS* semânticos. No domínio da informação geográfica, a GML aparece como um modelo padrão para geometria de elementos geográficos, mas não provê um padrão para a informação do domínio. Logo, a GML somente disponibiliza recursos para definir a estrutura de um documento válido. Falta, ainda, dispor a semântica dos elementos no documento para que pessoas e máquinas possam interpretá-los. Entrada bibliográfica.

Com a formalização do uso de ontologias pelos *WebGIS*, pode-se ter sistemas que permitam tanto consultas por usuários (pessoas) quanto por sistemas automatizados (agentes), permitindo, dessa forma, que agentes façam boa parte do trabalho de consulta e análise de informações contidas nos *WebGIS*. Como informações geográficas relacionam-se com vários domínios, existem diferentes definições, conceitos e relações para um mesmo objeto. Entrada bibliográfica. Por exemplo, os dados populacionais de um setor de estatísticas populacionais são diferentes em termos de atributos do que os mesmos dados em um departamento de vendas. Quando um departamento necessita de dados do outro, é necessário que os dados sejam transformados. A combinação de uma ontologia de domínio e a GML pode resolver este problema, possibilitando a interoperabilidade semântica para os *WebGIS* de uma forma automatizada e inteligente.

Como exemplo das possibilidades da *Web Semântica* e dos sistemas *WebGIS* na melhoria da busca de informações. Entrada bibliográfica, tem-se a situação em que um usuário necessita obter as informações sobre determinados lagos em uma região. Para isso, utiliza um mecanismo tradicional de busca na *Web* o remete a resultados que em geral contém apenas as palavras chaves informadas, sem nenhuma verificação se as mesmas estão sendo usadas da forma (conceito) que se deseja. Com o uso de ontologias este problema seria resolvido, pois far-se-ia uma consulta já especificada dentro de conceitos e elementos de uma ontologia, melhorando tanto a filtragem dos resultados

quanto permitindo consultas mais profundas, graças aos recursos de inferências permissíveis por agentes através das ontologias.

4. Trabalhos relacionados

Em relação aos *WebGIS*, pode-se citar alguns trabalhos que analisam a sua integração e sua evolução para *WebGIS Semânticos* (seja pela *Web Semântica* ou através de outros meios).

O sistema *OntoCarta* é uma aplicação que visa facilitar navegação em mapas pelo uso de ontologias. Permite uma navegação onde o usuário não precisa achar elementos no mapa através da manipulação direta deste, mas pode fazê-lo através de uma busca conceitual, utilizando uma ontologia. Desta forma, pode-se localizar “São Paulo”, indicando que se trata da *Cidade* e não do *Estado* Entrada bibliográfica. A ontologia usada é apresenta na figura Ilustração.

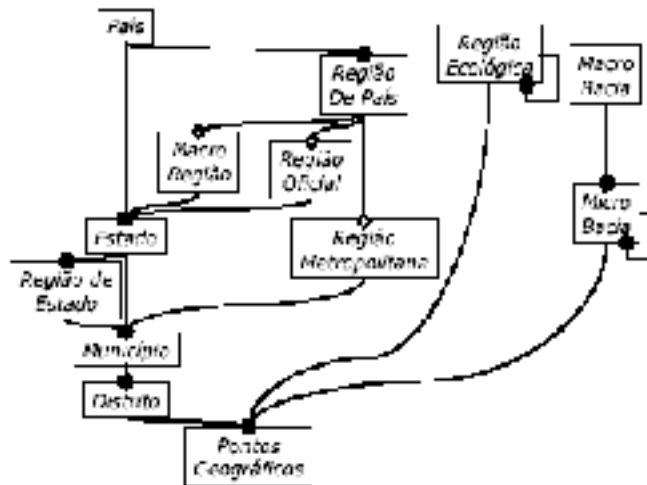


FIGURA Intervalo numérico Ilustração - Ontologia Utilizada pelo OntoCarta

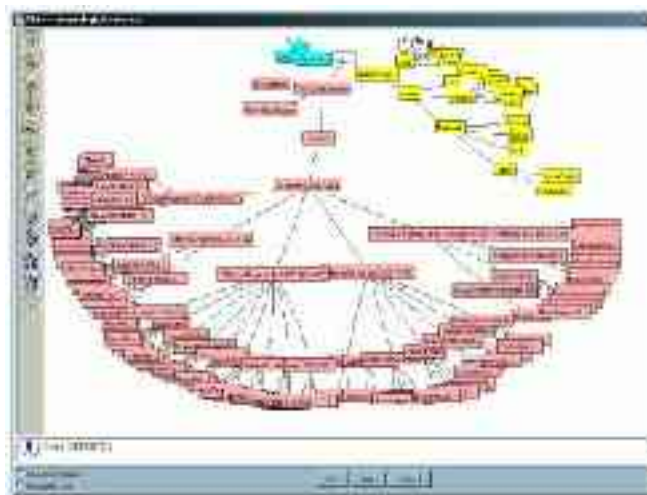


FIGURA Intervalo numérico Ilustração - Exemplo da árvore hiperbólica utilizada pelo OntoCarta

Para a apresentação dos dados são utilizados desenhos em formato SVG e para a navegação nos mapas é utilizada uma árvore hiperbólica, que representa os conceitos do

mapa de tal forma que os conceitos mais próximos do *conceito central*, no momento, são mais visíveis, enquanto conceitos mais distantes ficam em segundo plano, facilitando a navegação (figura Ilustração).

Sobre a integração entre SIGs, o artigo *Semantic and Interoperable WebGIS* Entrada bibliográfica faz uma análise da evolução da *Web* (referente a evolução trazida pela *Web Semântica*) e como esta pode beneficiar os *WebGIS*. Seu foco principal é a demonstração de como o uso de uma ontologia para auxiliar os *WebGIS* facilita a integração das informações e permite consultas mais complexas.

Uma idéia de arquitetura para os *WebGIS Semânticos* é apresentada na figura Ilustração Entrada bibliográfica. Nesta, têm-se vários *Clientes* de diferentes domínios, com consultas e ontologias específicas. Um *Mediator* executa uma tradução das consultas para as ontologias das fontes de dados (sistemas GIS) disponíveis, realizando consultas em uma ou mais bases de dados através das ontologias e do uso de regras de inferência (possível pelo uso de ontologias). Os dados obtidos são traduzidos para a ontologia dos *Clientes* e retornados.

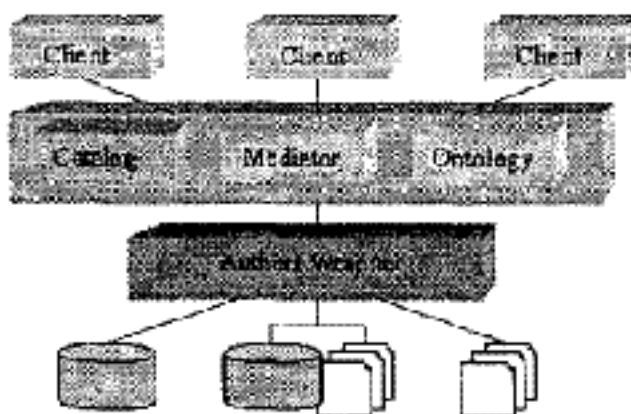


FIGURA Intervalo numérico Ilustração - A arquitetura proposta

Da mesma forma, outros trabalhos introduzem a idéia do uso de semântica para sistemas de *WebGIS*. Um aspecto interessante é a introdução de quatro formas básicas de um sistema na *Web* apresentar ao usuário informações com semântica Entrada bibliográfica:

- *Linguagem natural com marcação mínima* - Representa o que se tem atualmente com HTML ou XHTML. Não pode ser aproveitado por agentes automatizados de pesquisa, sendo útil apenas para usuários finais (pessoas);
- *Metadados simples* - Representados por linguagens de marcação como o XML, permitem uma representação semi-estruturada das informações. A semântica das informações fica confinada ao sistema de pesquisa e a interface, enquanto que os dados armazenados não possuem semântica (significado) associadas;
- *Modelos de dados* - Representados basicamente por RDF, ou seja, documentos que possuem uma estrutura conceitual conhecida. Através destes é possível estabelecer consultas envolvendo os conceitos dos dados armazenados sem conhecimento prévio da estrutura dos mesmos;
- *Lógica (modelo-teórico)* - Provê uma correspondência entre termos e entidades do mundo real, de forma a permitir uma automatização do processo de consulta

e cruzamento de dados. Ela é composta basicamente de interfaces, documentos e sistemas de pesquisa

Estas formas de apresentação de dados são equivalentes à evolução que a *Web Semântica* trouxe para a *Web*, proporcionando significado aos dados. Dois aspectos básicos devem ser garantidos para que se possa ter sistemas *WebGIS* integrados e com efetiva semântica das informações armazenadas Entrada bibliográfica:

- Uma forma padrão de consultar sistemas de *WebGIS*;
- Métodos para determinar as fontes (sistemas *WebGIS*) disponíveis adequadas à execução de uma determinada consulta.

Uma outra perspectiva vê que o crescimento da *Web* torna disponível fontes de dados ricas em muitas áreas Entrada bibliográfica. Algumas destas fontes armazenam fatos diariamente, os quais muitas vezes envolvem descrições geográficas. Estas descrições podem ser percebidas como dados georeferenciados indiretamente como, por exemplo, endereços, números de telefones, códigos postais e nomes de lugares. Nesta linha, a *Web* torna-se um grande banco de dados geo-espacial, provendo informação local e regional atualizada. Assim, a *Web* pode ser utilizada como uma importante fonte de informações geográficas e a extração de dados georeferenciados indiretamente pode alimentar aplicações de *WebGIS*. Com isso, é possível implementar SIGs urbanos que, por exemplo, usam endereços postais como base para integração de dados de diferentes fontes na *Web*, combinando estes dados com imagens de alta definição.

5. *WebGIS* atuais

Sistemas de *WebGIS* já existem e estão em uso na *Web*. Porém, em geral, apenas disponibilizam uma interface para consulta de informações, em alguns casos, uma interface para exportação de dados. Não se percebe, ainda, uma forte intenção de permitir a integração e consulta automatizada a estes sistemas por parte de seus desenvolvedores.

O OGC (*Open Geospatial Consortium* – <http://www.opengeospatial.org>) busca estabelecer padrões para serviços geoespaciais e de localização. Através destes padrões, pretende viabilizar uma melhor integração e acesso aos sistemas de *WebGIS* na Internet. Porém, o OGC ainda não faz uso de recursos da *Web Semântica* (como ontologias). Como exemplo de um produto destes padrões, tem-se o *Mapbender* (<http://www.mapbender.org/>), um *framework* para acesso a *WebGIS* compatíveis com os padrões do OGC.

O *site WebGIS* (<http://www.webgis.com>) contém um grande conjunto de mapas para utilização em SIGs, porém, não disponibiliza uma interface de pesquisa ou uma forma de integrar suas informações com outros sistemas, apresentando apenas uma navegação manual para a obtenção dos mapas desejados (que devem ser recuperados e acessados localmente).

Em *WebGIS.net* (<http://www.webgis.net>) destacam-se recursos mais avançados como migração de mapas, projeto de bancos de dados, transações de mapeamento *on-line* etc., além de recursos para hospedar sistemas de *WebGIS* de terceiros. Mas, ainda não possui recursos para facilitar o acesso às informações através do uso de semântica (como ontologias).

Uma aplicação para monitorar as reservas de determinados tipos de árvores na Califórnia, EUA, chamada *OakMapper* (<http://kellylab.berkeley.edu/SODmonitoring/>)

OakMapper.htm), permite realizar consultas sobre os dados nela armazenados. Em relação aos dois sistemas apresentados anteriormente, este último avança ao permitir a realização de consultas pelo usuário sobre os dados do sistema (tanto consultas textuais quanto consultas através de ferramentas gráficas). Por outro lado, esta aplicação não utiliza aspectos semânticos como ontologias para montar as consultas, assim como não disponibiliza uma interface acessível por agentes para consultas automatizadas e nem possui recursos para receber ou exportar informações.

Recentemente, tanto o *Google* (<http://local.google.com> e <http://maps.google.com>) e o *Yahoo!* (<http://local.yahoo.com>), ambos *sites* de pesquisa, introduziram ferramentas que permitem a visualização e manipulação de mapas (*Google Maps*) e a busca de locais (*Google Local* e *Yahoo! Local*) (figura Ilustração). No caso dos recursos disponibilizados pelo *Google*, uma API (*Application Interface*) está disponível para ambos os serviços, permitindo que agentes façam consultas sobre os mesmos. Contudo, nos três serviços citados não existe ainda o uso de ontologias ou recursos mais avançados relacionados a semântica das informações. As consultas são simples e orientadas para a utilização por um usuário final (pessoa).



Figura Intervalo numérico Ilustração – Exemplo de consulta no Google Local

O OGC também lançou recentemente o *Geospatial Semantic Web (GSW) Interoperability Experiment (IE)* Entrada bibliográfica. Este projeto desenvolverá importantes atividades para a expansão da *Web Semântica Geoespacial*, onde a descoberta, consulta e consumo de conteúdo geoespacial são baseados em especificações semânticas formais. O GSW habilitará o significado de consultas geoespaciais para serem facilmente compartilhadas entre diferentes *softwares* e serviços *on-line*. Vale lembrar, muitas das peças do quebra-cabeça do GSW têm sido desenvolvidas nos últimos anos. Este experimento aponta para o aumento de especificações de mapeamento focadas na *Web* (como o *Web Feature Service* and *Filter Encoding*) com recursos de consulta semântica, através da definição de uma ontologia (uma estrutura hierárquica de conhecimento) para a comunidade de inteligência geoespacial. O experimento deverá

explorar uma arquitetura distribuída apropriada para suportar o uso de cenários específicos.

6. Considerações finais

A constante evolução da Internet tem mostrado que os métodos atualmente disponíveis para encontrar e usar informações na *Web* são, muitas vezes, insuficientes. Neste sentido, percebe-se que a transição da *Web* de um simples repositório de dados para uma fonte de informação depende de métodos de organização da informação totalmente novos. A *Web Semântica* surge, então, como promessa de obtenção de melhores métodos de recuperação da informação. Para isso, incorpora a semântica dos dados e explora esta semântica durante o processo de busca Entrada bibliográfica.

No futuro, espera-se que os *WebGIS*, com os conceitos e recursos da *Web Semântica*, permitam consultas que retornem as informações desejadas sem a necessidade de verificar resultados sem relação com o que foi requisitado. Consultas envolvendo vários *WebGIS*, com cruzamento e análise de informações automatizados, retornando as informações requisitadas (e não apenas dados) são, também, objetivos dos *WebGIS* e da *Web Semântica*. Entidades como o OGC trabalham neste sentido e começam a apresentar resultados como, por exemplo, o *Mapbender*, apresentado neste trabalho.

Referências bibliográficas

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American**. 2001.

BORGES, K. A. V. *et al.* The Web as a Data Source for Spatial Databases. *In*: SIMPÓSIO BRASILEIRO DE GEOINFORMÁTICA, 5. **Anais...** Campos do Jordão: SBC/GeoINFO, 2003.

CÂMARA, G. *et al.* **Bancos de Dados Geográficos**. Curitiba: MundoGEO, 2005.

EGENHOFER, M. J. Toward the Semantic Geospatial Web. *In*: ACM INTERNATIONAL SYMPOSIUM ON ADVANCES IN GEOGRAPHIC INFORMATION SYSTEMS, 10. **Proceedings...** Virginia (EUA): ACM, 2002.

NOY, N. F. *et al.* Creating Semantic Web Contents with Protege-2000. *In*: **IEEE Intelligent Systems**, v. 16, Mar. 2001. p. 60-71.

OGC. **Geography Markup Language (GML) Implementation Specification 3.0**. OpenGeospatial, 2003. Disponível em: <<http://www.opengis.org>>.

OGC. **OGC to begin Geospatial Semantic Web Interoperability Experiment**. Press Release, 12/04/2005. Disponível em: <<http://opengeospatial.org/press/?page=pressrelease&year=0&prid=222>>.

SHANZHEN, Y. *et al.* Semantic and Interoperable WebGIS. *In*: INTERNATIONAL CONFERENCE ON WEB INFORMATION SYSTEMS ENGINEERING - WISE, 2. **Proceedings...** IEEE, 2001.

VENANCIO, L. R.; FILETO, R.; MEDEIROS, C. B. Aplicando Ontologias de Objetos Geográficos para Facilitar Navegação em GIS. *In*: SIMPÓSIO BRASILEIRO DE GEOINFORMÁTICA, 5. **Anais...** Campos do Jordão: SBC/GeoINFO, 2003.

VISSER, U.; STUCKENSCHMIDT, H. Interoperability in GIS - Enabling Technologies.
In: AGILE CONFERENCE ON GEOGRAPHIC INFORMATION SCIENCE, 5.
Proceedings... 2002. p. 291-297