

LOCUS: UM LOCALIZADOR ESPACIAL URBANO

Ligiane A. de Souza¹, Tiago M. Delboni¹, Karla A. V. Borges^{1,2}, Clodoveu A. Davis Jr.^{1,2} e Alberto H. F. Laender¹

¹UFMG – Universidade Federal de Minas Gerais – Departamento de Ciência da Computação – Av. Presidente Antônio Carlos, 6627, 31270-901, Belo Horizonte, MG, Brasil,

²PRODABEL – Empresa de Informática e Informação do Município de Belo Horizonte – Av. Presidente Carlos Luz, 1275, 31230-000, Belo Horizonte, MG, Brasil

Resumo: Este artigo apresenta o projeto do localizador espacial urbano Locus, cuja proposta é a implementação de um gazetteer associado a um mecanismo de busca e navegação. Quando em operação, o Locus será capaz de processar diferentes tipos de entrada alfanumérica, admitindo pesquisa por nome e endereço. O objetivo é executar sempre uma localização tão precisa quanto os dados disponíveis no gazetteer permitirem. Pretende-se trabalhar especificamente com o espaço urbano. Para isso, a ferramenta está sendo fundamentada em uma ontologia de lugares urbanos. Em sua versão inicial, o Locus conterà dados de todo o Brasil, pelo menos ao nível de divisão municipal, e de Belo Horizonte, chegando ao nível de endereço urbano individual.

Palavras-chave: gazetteer; georreferenciamento; ontologia espacial.

1. INTRODUÇÃO

Quando queremos indicar a localização de um lugar, quase sempre o fazemos pelo seu nome ou endereço. Este tipo de referência, onde a localização não aparece explicitamente na forma de coordenadas espaciais, é conhecido como referência espacial indireta (Borges et al., 2003). Muito comuns em nosso discurso falado e escrito, essas referências podem se transformar em informação preciosa quando formalmente representadas em

um SIG, isto é, quando convertidas em referências diretas na forma de um conjunto de coordenadas geográficas. A geocodificação trata exatamente deste problema: a identificação de uma posição na superfície terrestre a partir de um endereço alfanumérico (Davis et al., 2003). A proposta do trabalho apresentado neste artigo é a implementação do Locus, um sistema cujo objetivo principal é a localização de pontos de interesse a partir de endereços ou de referências a lugares urbanos. Trata-se de uma ferramenta que materializa conceitos de raciocínio espacial qualitativo (*qualitative spatial reasoning*) (Frank, 1996) e ainda permite inferir outros relacionamentos espaciais entre o ponto de interesse e sua vizinhança.

Um elemento fundamental para o sucesso do projeto é a construção de um *gazetteer* digital. Gazetteers são catálogos de nomes de lugares, onde cada lugar vem acompanhado de sua localização (Hill, 2000). Quando consultamos um atlas geográfico, geralmente utilizamos um *gazetteer* (o índice do atlas) para encontrar o que desejamos. Alguns *gazetteers* digitais podem ser encontrados atualmente na Internet (ADL, 2004; Getty Thesaurus, 2004). No Locus, o *gazetteer* será o componente responsável pela identificação do repositório onde os dados geográficos necessários poderão ser encontrados.

O Locus será capaz de reconhecer referências a lugares pelo nome, incluindo nomes não oficiais e antigos, e pelo endereço. No caso de busca por endereço, diferentes estratégias (Davis et al., 2003) serão empregadas com o objetivo de executar a localização da forma mais precisa possível. O *gazetteer* do Locus também irá conter e permitir consulta a outros tipos de referências espaciais indiretas, como códigos de CEP, DDD e prefixo telefônico.

A implementação do Locus faz parte de uma linha de pesquisa mais ampla, cujo objetivo é compreender a forma como as pessoas se referem a lugares e situações espaciais, para utilizar este conhecimento na identificação de contexto geográfico em páginas Web. Uma das etapas iniciais desta pesquisa foi a especificação de uma ontologia de lugares, onde estão formalmente definidos os conceitos presentes no espaço geográfico no domínio específico das cidades. A modelagem conceitual do banco de dados geográfico do Locus baseou-se nesta ontologia. O objetivo é que a interface do localizador reflita esses conceitos, permitindo a navegação do usuário pela ontologia.

O artigo está estruturado da seguinte forma: a Seção 2 apresenta a *ontologia de lugares urbanos*, que fundamenta o projeto. A Seção 3 descreve o modelo conceitual do *gazetteer*. A Seção 4 apresenta a arquitetura do sistema. A Seção 5 mostra os tipos de consulta possíveis, incluindo alguns exemplos. Finalmente, a Seção 6 contém alguns detalhes da implementação e do estágio atual em que se encontra o desenvolvimento do Locus.

2. ONTOLOGIA DE LUGAR

Uma ontologia descreve uma certa realidade com um vocabulário específico, usando um conjunto de premissas de acordo com o sentido intencional das palavras do vocabulário. Apesar da importância que ontologias vêm adquirindo para representar aspectos da informação, há poucos estudos sobre representação prática de conceitos de lugar, especialmente com o propósito de recuperação de informação. Nesta direção existem os trabalhos de Alani et al. (2001) e o Projeto Spirit (Jones et al., 2002).

Uma ontologia do espaço geográfico difere de outras porque a topologia e os relacionamentos “todo-parte” assumem um papel de destaque no domínio geográfico. A *ontologia de lugar* (Figura 1) é uma ontologia particular do espaço geográfico, definida como um conjunto de conceitos dentro de um domínio particular, no caso restrita a locais urbanos. Essa ontologia descreve feições naturais, objetos ou lugares que possuem significado para uma comunidade urbana, incluindo os relacionamentos entre eles. Um lugar é uma descrição de aspectos do espaço e possui identidade própria. O lugar é mais que uma geometria ou uma topologia, ele inclui um aspecto cognitivo e reflete como as pessoas percebem e usam a informação geográfica.

A ontologia explora a estrutura hierárquica do espaço, onde regiões são subdivididas em outras regiões (subdivisão territorial). Neste domínio é possível inferir os relacionamentos espaciais, topológicos, de continência e adjacência, utilizando apenas o conhecimento geográfico da estrutura hierárquica. Além da hierarquia de subdivisão territorial, a ontologia ainda inclui conceitos sobre o sistema de endereçamento postal e sobre os locais que constituem pontos de referência para a população.

No Locus, a ontologia é utilizada para dois propósitos: ela serviu de base para a modelagem conceitual (seção 3) e é também utilizada para contextualizar o resultado das consultas (seção 5).

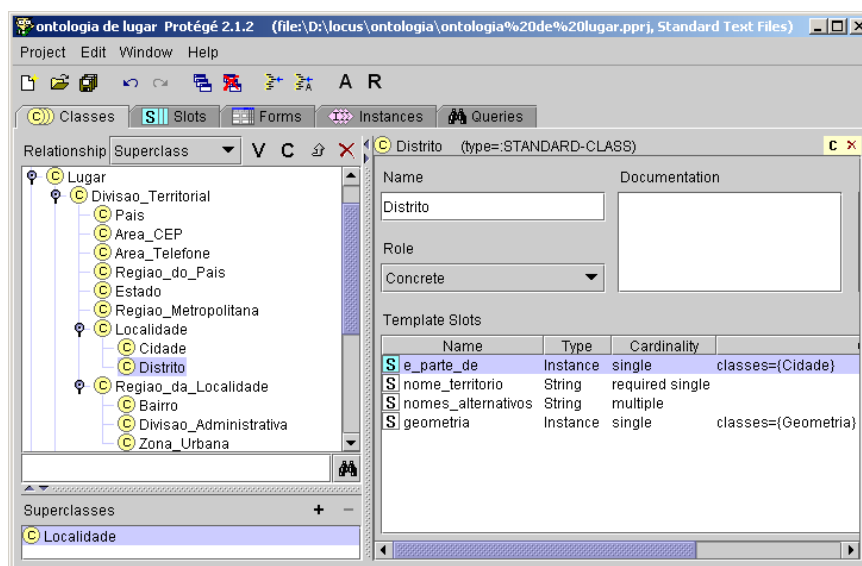


Figura 1. Ontologia de lugar descrita no Protégé

3. MODELAGEM CONCEITUAL

O modelo OMT-G (Borges et al., 2001) foi utilizado na modelagem conceitual do Locus. O esquema do gazetteer é apresentado na Figura 2. Nele, observamos que a classe *Lugar* é a superclasse de todas as entidades espaciais presentes no gazetteer. Toda instância desta classe recebe um identificador único, um valor indicativo do grau de certeza sobre a localização do dado espacial (Davis et al., 2003) e um nome (com exceção apenas das instâncias da classe *Endereço*). Se preciso, nomes alternativos ou anteriores também podem ser utilizados para identificar um local. O esquema prevê que várias relações espaciais podem ocorrer entre dois lugares. Estas relações podem ser de quatro tipos distintos: topológicas (*dentro de*), métricas (*a 100 metros de*), de ordem (*em frente de*, *ao lado de*) ou fuzzy (*perto de*) (Egenhofer e Franzosa, 1991; Pullar e Egenhofer, 1998). Um conjunto de relações estará disponível na interface do Locus para a elaboração de consultas (vide Seção 5).

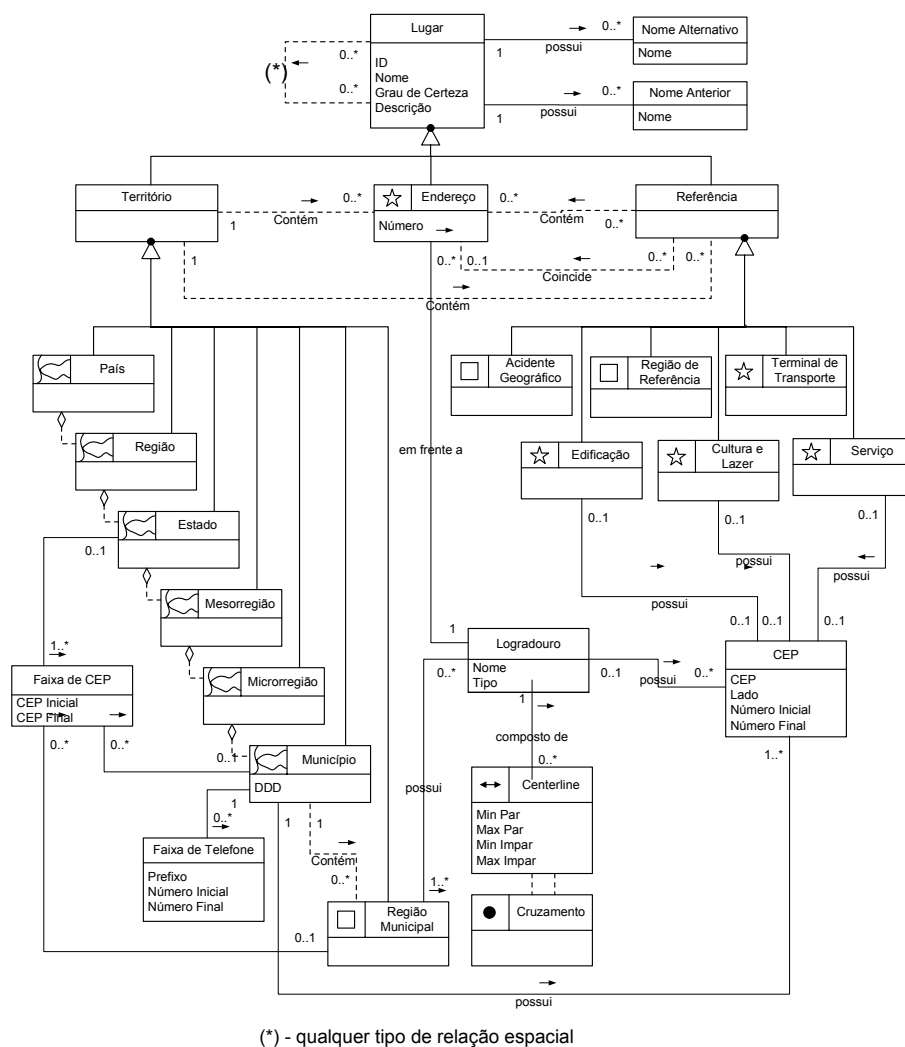


Figura 2. Esquema conceitual OMT-G do gazetteer.

A classe Lugar se especializa em três classes: Território, Endereço e Referência. Os territórios representam as divisões político-administrativas. A estrutura hierárquica das subdivisões dos territórios aparece sob a forma de agregações espaciais. A exceção fica por conta da classe Região Municipal, que pode representar distritos, bairros, divisões administrativas ou zonas urbanas. Logo, a associação entre Município e Região Municipal é de simples continência. Faixas de CEP podem ser atribuídas a estados e municípios. E prefixos e seqüências de números telefônicos podem ser associados aos municípios.

Os objetos da classe *Endereço* correspondem aos endereços urbanos com localização conhecida. Estes objetos são armazenados individualmente (um ponto para cada endereço). A opção pelo registro individual simplifica a representação e consegue tratar adequadamente problemas que são comuns em cidades brasileiras, como logradouros que possuem numeração irregular.

Lugares conhecidos da população e utilizados como pontos de referência são representados na classe *Referência*. Esta classe se especializa em classes mais representativas do tipo do ponto de referência, como *Acidente Geográfico*, *Terminal de Transporte*, *Cultura* e *Lazer* etc.

Por fim, a classe *CEP* trata daqueles lugares para os quais os Correios atribuem um código de endereçamento único. Isto ocorre para três tipos de lugar: municípios, logradouros de grandes municípios e alguns pontos de referência (grandes usuários dos Correios, como edifícios, universidades etc.). No caso dos logradouros, há aqueles para os quais os Correios atribuem mais de um CEP, seccionando o logradouro por faixas de numeração ou pelo lado (lado par e lado ímpar).

Os dados que estarão disponíveis na versão inicial do Locus são provenientes principalmente do IBGE (malha municipal do ano de 2001), dos Correios, da ANATEL – Agência Nacional de Telecomunicações e da Prodabel – Empresa de Informática e Informação do Município de Belo Horizonte.

4. ARQUITETURA DO SISTEMA

A Figura 3 apresenta um diagrama da arquitetura do sistema, destacando seus componentes principais e a interação entre estes durante uma consulta. A dinâmica destas consultas é detalhada na próxima seção. Como pode ser observado no diagrama, são admitidos dois tipos de entrada: endereços e nomes de lugares. Endereços sempre passam por uma estruturação (*parse*) antes da sua busca (*match*) no gazetteer. O resultado da fase de match é uma lista de candidatos, que passa por uma classificação (*ranking*), seguido da seleção pelo usuário. Por fim, a ontologia de lugar é utilizada para contextualizar o resultado selecionado, permitindo, através de uma navegação interativa o acesso a informações adicionais a respeito do local.

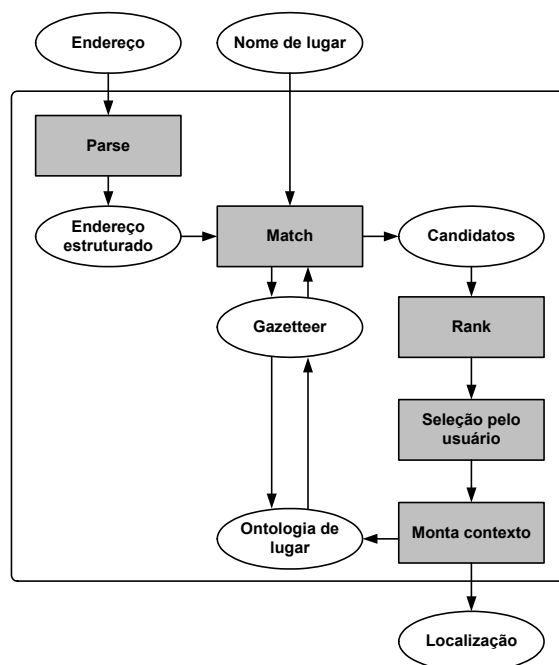


Figura 3. Arquitetura do Sistema.

5. CONSULTAS

Dois tipos de consulta estarão disponíveis: uma simples e outra avançada. O diagrama apresentado na Figura 3 corresponde a uma consulta do tipo simples, que pode se dividir em duas etapas: pesquisa e seleção. A etapa de pesquisa corresponde à entrada do nome do lugar ou do endereço que se deseja consultar (Figura 4).

Consulta Simples

Santa Helena	Pesquisar
---	--

Figura 4. Tela de consulta simples.

No caso da entrada de um nome de lugar, o Locus executa uma busca no gazetteer e retorna todas as entidades cujo nome é igual ou semelhante à entrada. Um processo de classificação das entidades é aplicado, criando um *ranking*. Diferentes critérios são utilizados para criação deste *ranking*, entre

eles o tipo da entidade e a semelhança do nome do lugar com o digitado pelo usuário. O resultado é uma lista de lugares, ordenada decrescentemente pelo *ranking*, que é retornada para o usuário, que então executa a seleção do resultado que mais lhe convém (Figura 5).

Seleção de Resultados

Resultados da pesquisa: "Santa Helena"

- 1) [Santa Helena](#) (Município de [Paraíba](#))
- 2) [Santa Helena](#) (Município de [Maranhão](#))
- 3) [Santa Helena](#) (Município de [Santa Catarina](#))
- 4) [Santa Helena](#) (Município de [Paraná](#))
- 5) [Santa Helena](#) (Bairro de [Belo Horizonte](#), [Minas Gerais](#))
- 6) [Rua Santa Helena](#) (Logradouro de [Belo Horizonte](#), [Minas Gerais](#))

Figura 5. Tela de seleção de resultados.

O resultado final é então exibido (Figura 6). É importante mencionar que o *layout* da tela de resultado e os *links* de navegação são montados dinamicamente e dependem do tipo do lugar selecionado, que, em nosso exemplo, é um município (Santa Helena, Paraná). Isto significa que, a partir desta tela, é possível consultar o estado ao qual o município pertence, sua região, municípios adjacentes, regiões municipais etc. Estas opções possibilitam uma navegação interativa pelos conceitos da *ontologia de lugares*.

Resultado

Santa Helena

Município

Estado: [Paraná](#)

Mesorregião: [Oeste Paranaense](#)

Microrregião: [Toledo](#)

Municípios adjacentes:

Regiões municipais:

DDD: [45](#)

CEP: 85892-000

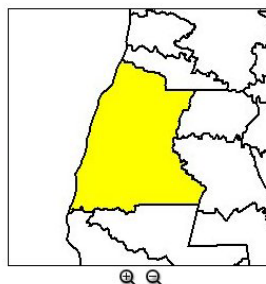


Figura 6. Tela de resultado da consulta.

Este exemplo mostra a consulta a um território. A consulta por pontos de referência ocorre da mesma maneira. Já na consulta por endereço, o usuário é induzido a digitar o mesmo em um formato pré-definido. Conceitualmente, a localização espacial de endereços (geocodificação) é um processo

constituído por três etapas: o tratamento do endereço alfanumérico semi-estruturado (*parsing*), o estabelecimento de uma correspondência entre o endereço estruturado e o banco de dados (*matching*) e a atribuição das coordenadas geográficas ao evento que está sendo alvo da geocodificação (*locating*) (Davis et al., 2003). No Locus, a imposição da entrada do endereço em um formato pré-estabelecido diminuirá a incidência de erros na fase de parsing. Solicitada a consulta, o algoritmo de matching entra em ação, executando uma sequência de tentativas de localização do endereço, partindo da possibilidade de geocodificação mais precisa (endereço individual armazenado no gazetteer) e avançando para aproximações baseadas nos dados disponíveis (interpolação pela faixa de numeração da *centerline*, centróide do bairro, coordenada aleatória dentro do município etc.).

Já na consulta avançada, é possível utilizar relações espaciais para filtrar o resultado da pesquisa (Figuras 7 e 8). Nesta opção, o usuário deve entrar com uma trinca de valores: um ponto de interesse (*o que* se deseja encontrar), uma relação espacial e um ponto de referência (*onde* se deseja encontrar). Isto permite consultas mais elaboradas, como “ponto de ônibus” “perto da” “Praça da Liberdade”, “lanchonete” “a 1000 metros da” “Rua São Paulo, 100” e “McDonalds” “perto da” “Praça Carlos Chagas”. Como pode-se perceber nestes exemplos, o ponto de interesse pode ser o nome ou um tipo de lugar.

Consulta Avançada

o que

perto de

onde

Figura 7. Tela de consulta com relação fuzzy.

Consulta Avançada

o que

metros de

onde

Figura 8. Tela de consulta com relação métrica.

Tal como na consulta simples, o usuário é solicitado a selecionar o ponto de referência e, se necessário, o ponto de interesse. A seguir, a relação espacial é interpretada e uma operação espacial é elaborada. Relações de ordem ou *fuzzy* são interpretadas levando-se em consideração a natureza dos pontos de interesse e de referência. Isto significa que valores quantitativos são aplicados a relações espaciais qualitativas. Para o exemplo da Figura 7, uma área de influência de 500 metros em torno da Praça da Liberdade, Belo Horizonte, pode ser utilizada para localizar os pontos de ônibus “perto da” praça (Figura 9).

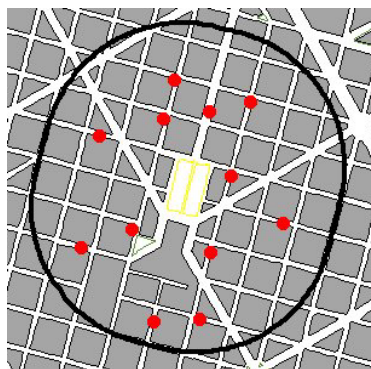


Figura 9. Resultado da consulta espacial.

6. IMPLEMENTAÇÃO E ESTÁGIO ATUAL

O Locus é totalmente baseado em software livre. A interface Web do sistema está sendo desenvolvida em JSP. O SGBD PostgreSQL, acrescido do módulo PostGIS, foi a escolha para armazenamento de dados espaciais. Para a criação e exibição dos mapas via Web, será utilizado o servidor de mapas MapServer. Todos os softwares rodam em plataforma Linux.

Os primeiros meses de desenvolvimento do sistema foram dedicados à modelagem, coleta e estruturação dos dados. No estágio atual, o banco de dados do Locus conta com as divisões político-administrativas de todo o país até o nível de município, incluindo CEPs, DDDs e prefixos telefônicos, além de logradouros, bairros, referências e endereços de Belo Horizonte. Serão ainda incorporados ao banco de dados catálogos de logradouros e bairros de outras cidades brasileiras. A etapa seguinte compreende o projeto da interface e a implementação dos algoritmos de busca e geocodificação.

7. CONCLUSÕES

Este artigo apresentou o Locus, um sistema para a localização e visualização de pontos de interesse (endereços ou referências). O projeto baseia-se em uma ontologia de lugar. Esta ontologia foi utilizada na modelagem do sistema e na definição dos processos de busca e apresentação do resultado das consultas. Um gazetteer é empregado para o armazenamento das instâncias da ontologia. Nas consultas ao sistema, o objetivo é o de recuperar sempre a localização mais precisa que os dados disponíveis no gazetteer permitirem. Quando totalmente implementado, o Locus possibilitará consultas simples e avançadas. Neste último caso, relações espaciais são utilizadas na definição da consulta.

Nossa expectativa é de que a ferramenta venha a ser empregada como componente de várias aplicações onde a localização de eventos e entidades através de endereços é necessária. Usos previstos da ferramenta incluem as áreas de saúde pública (em particular na epidemiologia urbana), criminalidade e segurança pública (Projeto SAUDAVEL, 2004), educação (Pinto, 1999) e planejamento urbano. Além disso, o Locus é essencial para apoiar a identificação de contexto geográfico em páginas Web (Borges et al., 2003) e no georreferenciamento de objetos de bibliotecas digitais (Barros e Laender, 2004).

8. REFERÊNCIAS

- Alani, H., Tudhope, D., Jones, C. B., 2001, Geographical information retrieval with ontologies of place, *Spatial Information Theory: Foundations of Geographic Information Science - COSIT 2001*. 322-335.
- ADL Gazetteer, 2004, (19 de outubro de 2004); <http://www.alexandria.ucsb.edu>.
- Barros, E. G., Laender A. H. F., 2004, Uma biblioteca digital para o Peld Brasil, *Resumos do Simpósio Internacional sobre Projetos Ecológicos de Longa Duração*. 57-59.
- Borges, K. A. V., Davis, Jr., C. A., Laender, A. H. F., 2001, OMT-G: An object-oriented data model for geographic applications, in *GeoInformatica*. 5(3): 221-260.
- Borges, K. A. V., Laender, A. H. F., Medeiros, C. B., Silva, A. S., Davis Jr., C. A., 2003, The Web as a data source for spatial databases, *V Simpósio Brasileiro de Geoinformática - GeoInfo 2003, Campos do Jordão (SP)*.
- Davis Jr., C. A., Fonseca F. T., Borges, K. A. V., 2003, A Flexible Addressing System for Approximate Geocoding. *V Simpósio Brasileiro de Geoinformática - GeoInfo 2003, Campos do Jordão (SP)*.
- Egenhofer, M. J., Franzosa, R. D., 1991, Point-set topological spatial relations. *International Journal of Geographical Information Systems*. 5(2): 161-174.
- Frank, A. U., 1996, Qualitative spatial reasoning: cardinal direction as an example, *International Journal of Geographical Information Systems* 10(3): 269-290.
- Getty Thesaurus, 2004, (19 de outubro de 2004); <http://www.getty.edu>.

- Hill, L.L., 2000, Core elements of digital gazetteers: placenames, categories, and footprints, *Proceedings of the 4th European Conference on Research and Advanced Technology for Digital Libraries*, 280-290.
- Jones, C. B., Purves, R., Ruas, A., Sanderson, M., Sester, M., van Kreveld, M., Weibel, R., 2002, Spatial information retrieval and geographical ontologies. An overview of the SPIRIT project, *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development Information Retrieval*, Tampere, Finland, 387-388.
- Pinto, M. V., 1999, Cadastramento escolar: democratização do acesso à escola pública. *Informática Pública*, 1(2): 139-156.
- Projeto SAUDAVEL, 2004, (19 de outubro de 2004); <http://saudavel.dpi.inpe.br/>.
- Pullar, D. V., Egenhofer, M. J., 1998, Towards the defaction and use of topological relations among spatial objects. *Proceedings of the 3rd International Symposium on Spatial Data Handling*, Columbus, Ohio, 225-24.