

Revista Brasileira de Cartografia (2013) N^o 65/6: 1127-1142
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto
ISSN: 1808-0936

USO DE ONTOLOGIAS PARA BUSCA DE DADOS GEOESPACIAIS: UMA FERRAMENTA SEMÂNTICA PARA A INFRAESTRUTURA NACIONAL DE DADOS ESPACIAIS

Using Ontologies to Search Geospatial Data: a Semantic Tool for the National Spatial Database Infrastructure

Silvana Philippi Camboim & Claudia Robbi Sluter

Universidade Federal do Paraná – UFPR

Curso de Pós-Graduação em Ciências Geodésicas

Caixa Postal 19001 - Centro Politécnico, Jardim das Américas - 81531-990 Curitiba - Paraná, Brasil

{silvanacamboim, robbi@ufpr.br}

Recebido em 27 de Novembro, 2012/Aceito em 03 de Maio, 2013

Received on November 27, 2012/ Accepted on May 03, 2013

RESUMO

A Infraestrutura Nacional de Dados Espaciais – INDE-BR tem como objetivo a padronização e o compartilhamento de informações geográficas no Brasil. Como em outras Infraestruturas de Dados Espaciais (IDEs) criadas mundialmente, é um desafio promover a utilização ampla e integrada de dados e serviços geográficos. Dada a multiplicidade de produtores e usuários envolvidos, a utilização de ontologias para representar o conhecimento empregado na modelagem dos dados se mostra uma ferramenta potencialmente útil. Elas podem apoiar a usabilidade em diversos níveis, tanto para compatibilizar terminologias entre grupos de usuários, quanto para promover a interoperabilidade entre sistemas e serviços disponíveis. Neste trabalho, é explorada uma aplicação de Ontologias, para realização de Buscas Semânticas, uma forma de facilitar o acesso dos usuários a dados distribuídos. É discutido como uma ferramenta como esta pode simplificar operações que exigem muitos passos para serem efetuadas usando as alternativas disponíveis atualmente.

Palavras chaves: Infraestrutura Nacional de Dados Espaciais, Busca Semântica, Ontologias.

ABSTRACT

The Brazilian National Spatial Database Infrastructure – INDE-BR aims to promote standardization and sharing of national geospatial information. As other Spatial Database Infrastructures (SDIs) created worldwide, it faces challenges to enable the wide and integrated use of geographical data and services. The use of Ontologies to represent the knowledge applied in data modeling is considered potentially useful because there are multiple data producers and users involved. They can support usability in several levels, both to reconcile terminology between users groups and to promote interoperability between systems and services available. This work explores one application of Ontologies, using them to enable Semantic Search, in a way to facilitate user access to distributed data. Additionally, there is a discussion about how this kind of tool can simplify operations that require many steps to execute using currently available alternatives.

Keywords: National Database Infrastructure, Semantic Search, Ontologies.

1. INTRODUÇÃO

Com a crescente disponibilidade de dados geoespaciais, tornaram-se relevantes os esforços para que essa informação seja estruturada para garantir a sua acessibilidade por parte de diversos grupos de usuários. Foi nesse contexto que foram criadas as Infraestruturas Nacionais de Dados Espaciais (INDEs), definidas como “*uma rede dinâmica composta de pessoas, dados, políticas, padrões e tecnologia para utilização melhorada de dados e serviços espaciais dentro da jurisdição de um país*” (GRUS et al., 2006). Desde o início dos anos 1990s foram registrados esforços para criação de tais redes. Em 2003, CROMPVOETS & BREGT identificaram mais de 120 países com INDEs em processo de estabelecimento, sendo as pioneiras a americana (1994) e a australiana (1996). Na América Latina, muitas INDEs já estão oficializadas e em funcionamento, como as do México, Chile e Argentina (ALBITES, 2003; IGLESIAS, 2008; RICKERT e MACHUCA, 2008).

No Brasil, a INDE (aqui, INDE-BR) foi criada em novembro de 2008 através do decreto 6.666/2009; que estabelece seus objetivos, dentre eles o de “*promover o adequado ordenamento na geração, no armazenamento, no acesso, no compartilhamento, na disseminação e no uso dos dados geoespaciais de origem federal, estadual, distrital e municipal, em proveito do desenvolvimento do País*”.

A INDE-BR é gerenciada pelo Comitê de Implantação da INDE (CINDE) na CONCAR - Comissão Nacional de Cartografia. Neste Comitê reúnem-se diversos representantes do governo, academia e indústria para executar o Plano de Ação para Implementação da INDE. Este plano foi lançado em abril de 2010, expondo detalhadamente as ações necessárias para que a INDE se realize, dividindo a execução em três ciclos que abrangem um horizonte de planejamento de 10 anos (CONCAR, 2010).

No âmbito da INDE-BR, a rede de servidores de dados capaz de reunir eletronicamente produtores, gestores e usuários dos dados geoespaciais envolvidos é denominada Diretório Brasileiro de Dados Geoespaciais, ou DBDG. Cada produtor de dados pode criar seu próprio servidor dentro do DBDG, constituindo um nó autônomo, ou então utilizar o nó central que é

gerenciado pelo IBGE (Instituto Brasileiro de Geografia e Estatística). Para inserção da instituição ao DBDG está prevista a assinatura de um Termo de Adesão, no qual constarão os dados, metadados e serviços a serem disponibilizados bem como o planejamento da manutenção do nó, caso seja executado pela própria instituição. O Portal Brasileiro de Dados Geoespaciais - SIG Brasil é a porta de acesso a estas informações quando disponibilizadas, através do qual os usuários podem explorar os recursos distribuídos do DBDG.

No capítulo 03 do Plano de Ação, foram identificados os principais atores da INDE-BR e os conjuntos de dados geoespaciais a cada um deles relacionados, de acordo com a tabela resumo apresentada a seguir:

Tabela 1: Exemplos de produtores de informação geoespacial no Brasil.

	Grupos de Informação Geoespacial	Exemplos de Produtores
Dados de Referência	Controle Geodésico	IBGE, Observatório Nacional.
	Cartografia Terrestre Básica (Mapeamento Geográfico, Topográfico e Cadastral).	IBGE, Prefeituras, Exército/DSG.
	Cartografia Especial (Mapeamento Náutico e Aeronáutico)	Aeronáutica/ICA e Marinha/DHN
	Subsidiários e Acessórios (Mosaicos Ortoretificados, Nomes Geográficos, Divisão Política-Administrativa, Terras Indígenas, Unidades de Conservação, Dados Fundiários, entre outros).	FUNAI, INCRA, ICMBio, IBGE, Órgãos Estaduais de Terra e Cartografia, Ministério das Relações Exteriores, INPE, entre outros.
Dados Temáticos	Meio Abiótico (Geomorfologia, Geologia, Solos, entre outros).	IBGE, Embrapa, DNPM, CPRM,
	Meio Biótico (Vegetação, Biomas, Biodiversidade, entre outros).	MMA e órgãos estaduais e municipais integrantes do SIS-NAMA
	Meio Antrópico (Zoneamento Ecológico-Econômico, Riscos, Agricultura, Transportes, entre outros).	Ministério da Agricultura, Transportes.

Com base no quadro é possível verificar a gama de atores envolvidos na produção de informação geoespacial no Brasil, mesmo considerando esta uma amostra não exaustiva

de instituições diversas, que possuem cultura, processos e objetivos próprios. Adicionalmente aos produtores, ainda mais diversos são os usuários da INDE-BR, definidos no Plano de Ação como sendo “os integrantes de órgãos federais, estaduais, municipais, distritais, dos órgãos da administração direta e indireta (autárquicas ou fundacionais); da iniciativa privada, de organizações civis, de organizações não governamentais (ONGs) e associações; e cidadãos (usuário comum que não está vinculado a nenhuma organização ou entidade), que utilizam os dados e metadados fornecidos pela INDE para integrá-los em diferentes aplicações, efetuando análises temáticas e de assuntos diversos, ou simplesmente utilizando os serviços do geoportais SIG Brasil” (CONCAR, 2010). Essa descrição retrata a diversidade dos usuários envolvidos, desde especialistas em distintas áreas de conhecimento até cidadãos em geral. Cada grupo destes possui seu vocabulário específico e nem sempre compartilha os mesmos conceitos usados pelos produtores da informação espacial quando nas etapas de criação e abstração da mesma, desde a realidade física até o ambiente computacional.

Um dos problemas mais comumente relatados por usuários de dados geoespaciais é a dificuldade de encontrar as informações relevantes para determinado uso. Um exemplo citado por BERNARD e CRAGLIA, 2005 é uma pesquisa da EC DG Environment em 50 empresas europeias que trabalham na preparação de Relatórios de Estudos de Impactos Ambientais (EIA), que obteve os seguintes resultados:

- Problemas para acessar dados existentes (70%),
- Dificuldade em achar qual dado está disponível (56%),
- O dado necessário não está disponível (51%),
- Dados de diferentes fornecedores não são compatíveis (47%),
- Dados existentes são de qualidade insuficiente (47%).

Portanto, se estabelece como uma questão importante para as INDEs o esforço em facilitar o acesso aos dados por parte dos usuários. Internacionalmente, é significativa a mudança de paradigma da primeira geração de IDEs criada na década de 1990, essencialmente voltada a dados e processos, para as atuais, mais orientadas

a serviços e preocupadas com as demandas dos usuários (BERNARD e CRAGLIA, 2005; FERDANDEZ e CASTELLANOS, 2006; LEITE JR et al., 2006 e MASSER, 2009). Para tanto, é necessário fortalecer a interação entre os diversos produtores de informação e os usuários, considerando os vários entendimentos da realidade envolvidos.

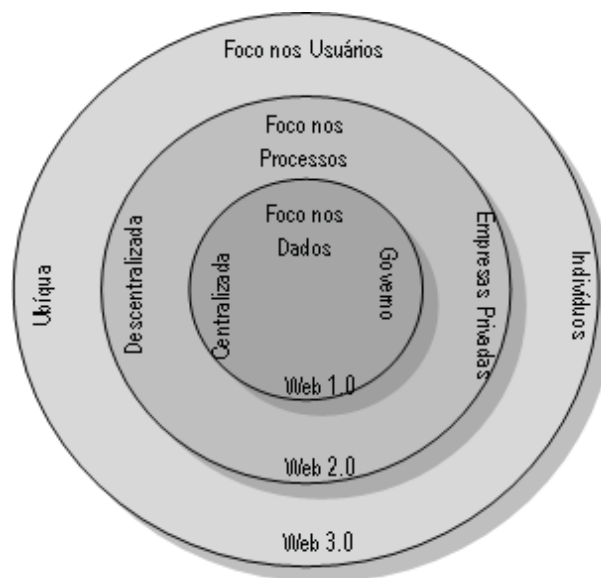


Fig. 1 - A evolução das IDEs.

A Figura 01 ilustra o desenvolvimento das IDEs, em paralelo com o desenvolvimento da Internet em geral. As diversas fases da evolução da Internet podem ser comparadas, em uma analogia, às permissões de sistemas de arquivo (PAN et al, 2009), sendo a primeira fase, ou web 1.0, onde seriam possíveis ações de “somente leitura” de conteúdo, a seguinte ou web 2.0, com a possibilidade de “leitura e escrita”, com o advento das redes sociais e participativas no estilo wiki, e finalmente a web 3.0, que permite a “leitura, escrita e execução”, numa alusão à inclusão de capacidades semânticas e à orientação a serviços. Igualmente, as IDEs em geral inicialmente estavam focadas em divulgar dados governamentais através de portais de dados geoespaciais centralizados. Com a expansão do número e quantidade de atores, incluindo produtores privados de dados, teve início uma nova fase centrada em processos e padrões criados em fóruns especializados como o OGC (Open Geospatial Consortium). Entretanto, segundo SADEGHI-NIARAKI et al., 2010 no desenho desses processos as preferências

e interesses dos usuários finais não foram considerados. Desta forma, surge atualmente uma nova fase, baseada numa participação mais ativa por parte do usuário e promovendo o papel do indivíduo como produtor e consumidor da informação geoespacial através de dispositivos móveis, cada vez mais populares. É a fase da interação social e do compartilhamento de informações em tempo real que tem se mostrado especialmente útil em situações que demandam respostas ágeis como desastres naturais e crises de emergência social.

No Brasil, a INDE-BR ainda está no seu ciclo inicial de implementação, fortemente identificado com as características de IDEs com foco nos dados. No entanto, com o crescimento da adoção de normas internacionais e produção de normas locais ou adaptadas, como o Perfil Brasileiro de Metadados e a Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV), a infraestrutura nacional está avançando para a segunda etapa do modelo. Os nós do DBDG espelham essa configuração de servidores distribuídos da IDE focada em processos. Nesse esquema de evolução das IDEs, uma fase não suplanta a anterior, ao contrário, ambas convivem com a extensão das capacidades inicialmente instaladas para novos horizontes de aplicação, o que aponta para o crescimento da preocupação com o usuário também na INDE-BR.

Dentre as propostas para esta terceira geração de IDEs estão as que buscam a solução de aproximação junto aos usuários utilizando ontologias (ou *Ontology-driven SDI* - SADEGHI-NIARAKI et al., 2010). Uma ontologia pode ser entendida como constituída por um vocabulário específico, usado para descrever uma certa realidade, mais um conjunto de pressupostos explícitos a respeito do significado pretendido para o vocabulário (GUARINO, 1998). A ontologia pode colaborar com a intermediação e descrição de conceitos para a comunicação entre humanos, no caso da utilização de jargão técnico de conhecimento exclusivo de grupos determinados de especialistas, ou entre humanos e computador, promovendo a melhora na usabilidade dos sistemas, e ainda entre computadores colaborando para a interoperabilidade entre sistemas.

Uma vez estabelecida pela comunidade de produtores e usuários da informação, a ontologia

pode ser utilizada para uma série de aplicações para construir novas formas de interação com a informação geoespacial. Uma dessas aplicações pode ser o auxílio à busca de informações no banco de dados ou serviços geoespaciais online. Seria possível então a criação de buscadores semânticos com o potencial de contextualizar as pesquisas e automatizar a geração de conhecimento através dos relacionamentos, ajudando assim a adaptar dentro dos vários contextos as intenções e linguagem de cada usuário.

Neste trabalho procurou-se descrever um buscador semântico que possa aprimorar a busca por dados espaciais no Diretório Brasileiro de Dados Espaciais, trazendo capacidades semânticas para a INDE brasileira sem exigir do usuário conhecimento de quais instituições são responsáveis pelos diversos tipos de informação, muito menos do modelo e da estrutura empregada para o armazenamento dos dados. Do contrário, em uma estrutura semântica, deve ser possível usar expressões coloquiais para que assim o sistema de buscas possa efetivamente auxiliar o acesso às informações.

2. BUSCA E RECUPERAÇÃO DE INFORMAÇÃO GEOESPACIAL

2.1 Busca de Informação Geoespacial

Uma característica básica de uma IDE é a facilitação do acesso e troca de informações geoespaciais. Para tanto, é fundamental a existência de ferramentas que possibilitem ao usuário interagir com as diversas bases de dados envolvidas, procurando por informações que atendam às suas necessidades, e para isso usando critérios espaciais, temporais ou outras características associadas às informações. Estas interações com as IDEs em busca de informação se dividem em duas frentes (KLIEN, 2005):

- **Descoberta de Informação Geoespacial:** numa estrutura de produtores e usuários distribuídos através da rede, a estrutura da IDE prevê a existência de catálogos de metadados onde os produtores possam publicar as características das informações disponíveis. Geralmente estes catálogos possuem ferramentas de buscas nos metadados para selecionar conjuntos de dados que atendam determinados critérios que estejam descritos pelos metadados (área de abrangência,

forma de execução, data de atualização, nível de detalhamento, entre outros). Para descrição do conteúdo da base de dados, são usadas palavras-chave, o que pode acabar gerando problemas na busca, caso seja usada terminologia na descrição do conteúdo diferente da empregada pelo usuário.

- Recuperação de Informação Geoespacial: uma vez descoberta uma fonte de dados geoespaciais na etapa anterior, existe a interação com o conjunto de dados em si, através da criação de consultas ou filtros das informações armazenadas. Este processo pode ser feito interagindo-se diretamente com Banco de Dados, como detalhado no item 2.2 a seguir, ou através de serviços online que serão descritos no item seguinte. Novamente, problemas semânticos podem ocorrer na interpretação do esquema conceitual usado quando da modelagem dos dados utilizados. Uma represa, por exemplo, pode ser entendida diferentemente em contextos distintos, como por exemplo no contexto de produção energética, de lazer ou como limite entre dois municípios.

2.2 Bancos de Dados Geográficos e Consultas SQL

Como estabelecido por CAMARA, 2005, o problema fundamental da Geoinformação é a “produção de representações computacionais do espaço geográfico”. O processo para essas representações é bem ilustrado pelo Paradigma dos Quatro Universos (Gomes e Velho), ilustrado na Figura 2:



Fig. 2 - Paradigma dos Quatro Universos.

O Universo Ontológico compreende os conceitos da realidade que serão representados computacionalmente, e, portanto é o primeiro nível de abstração. Em muitos Sistemas de Informação, esta etapa não é formalizada. O Universo Formal corresponde às abstrações formais necessárias, como modelos de dados e modelos conceituais. O Universo Estrutural é composto dos algoritmos necessários para implementação, a conversão e o armazenamento dos dados e seus relacionamentos topológicos.

Por fim, no Universo de Implementação, executa-se a conversão dos modelos anteriores em implementações computacionais.

Neste trabalho os dados estudados são de natureza espacial, portanto gerenciados em Bancos de Dados Geográficos. Para realizar os serviços de armazenamento, consulta e atualização de bancos de dados geográficos, foram criados os Sistemas de Gerência de Banco de Dados (SGBD) (FERREIRA et al, 2005). Os primeiros SGBD, desenvolvidos pela IBM com base no modelo proposto por E. F. Codd, em 1970, no artigo “A Relational Model of Data for Large Shared Data Banks”, são os chamados Bancos de Dados Relacionais. Eles são formados por registros que se relacionam entre si. Na década de 1980, ao mesmo tempo em que os Bancos de Dados Relacionais se consolidavam no mercado, novas aplicações levaram à criação dos SGBD Orientados a Objeto, que permitiam incorporar conceitos de orientação a objeto como polimorfismo e herança, e também abrindo as portas para o armazenamento de novos tipos de dados, como áudio, vídeo e dados geoespaciais (SILBERSCHATZ ET AL, 2005). Posteriormente, estes conceitos foram incorporados aos já então amplamente utilizados SGBD Relacionais, dando origem aos SGBD Objeto-relacionais, que são o padrão atual em termos de Bancos de Dados Geográficos. Foi essa evolução que permitiu a criação de extensões espaciais para SGBD não-espaciais já difundidos, como o Oracle Spatial e o PostGIS.

A origem do SQL (Structured Query Language ou Linguagem de Consulta Estruturada) foi contemporânea ao aparecimento dos SGBD Relacionais, nos anos 1970s. A linguagem chamada inicialmente de SEQUEL, recebeu posteriormente a denominação atual e foi adotada como padrão ISO no ano de 1987. A versão mais atual é da norma ISO/IEC 9075-1:2008, e é hoje usada pela maioria dos SGBD Relacionais e Objeto-Relacionais existentes. (FERREIRA et al, 2005).

Os comandos SQL são divididos em cinco subgrupos: a DML (Linguagem de Manipulação de Dados), usada para inserção, alteração e exclusão de dados; a DDL (Linguagem de Definição de Dados), para definir e alterar tabelas e outros elementos; a DCL (Linguagem de Controle de Dados), para criar e alterar

permissões de acesso aos dados; a DTL (Linguagem de Transação de Dados), para iniciar e manipular transações de dados; e a DQL (Linguagem de Consulta de Dados) que com apenas um comando, o SELECT, é a responsável pela criação de consultas, ou buscas, que são exploradas neste trabalho. Na tabela 02 está descrita a sintaxe do comando.

Tabela 2: Estrutura do comando select. Adaptado de Ferreira et al (2005).

Partes do Comando SELECT	
SELECT {*, coluna, expressões, funções,...}	Uma ou mais colunas, expressões aritméticas envolvendo as mesmas, ou funções definidas em SQL como, por exemplo, funções de agregação.
from {tabelas}	Tabelas das quais os dados serão selecionados
[where condição]	Condição a ser satisfeita para que os registros façam parte do resultado da busca
[group by colunas]	Agrupamento pela coluna especificada
[having condição]	Limita os grupos a serem mostrados àqueles que satisfaçam a condição
[order by colunas [asc desc]]	Ordenação a ser aplicada no resultado da busca

Por exemplo, o seguinte comando:

```
SELECT nome FROM município WHERE
populacao > 500.000 AND estado = "PR"
ORDER BY nome
```

retorna a lista de nome dos municípios, por ordem alfabética, cuja população seja maior que quinhentos mil habitantes no Estado do Paraná.

O SQL original não comporta dados e funções espaciais. Para isso a OGC (Open Geospatial Consortium) teve a iniciativa de criar uma norma específica para a área. A norma OpenGIS “*Geographic information — Simple*

feature access” (SFA), também homologada como ISO 19125 (OGC, 2006), é dividida em duas partes, sendo a primeira definidora da arquitetura e a segunda detalha a opção SQL. A norma se baseia em feições simples, formadas por entidades geométricas de ponto, curva e superfície em duas ou três dimensões. Na figura 03 estão definidas a hierarquia das classes de geometria possíveis. As feições são armazenadas em tabelas nas quais a parte não espacial é tratada pelas normas já definidas para o SQL, enquanto a parte geométrica é armazenada em colunas específicas cujas regras são definidas pela norma da OGC. A coluna espacial é associada a um sistema de referência espacial, de forma também descrita na primeira parte da norma, na qual são especificados atributos, métodos e operações com geometrias.

A segunda parte da norma ISO 19125 (OGC, 2010b), define um esquema SQL que suporta o armazenamento, a recuperação, a busca e a atualização de coleções de feições através da interface de nível de chamada SQL definida pela norma ISO/IEC 9075-3:2003. A partir dessa norma, critérios espaciais podem ser definidos nas cláusulas de busca SQL, podendo retornar geometrias nas buscas no lugar de resultados tabulares apenas. Alguns exemplos de métodos espaciais são: ST_Disjoint, ST_Intersects e ST_Touches, para as operações topológicas Disjuntas, Intersecção e Toca, respectivamente.

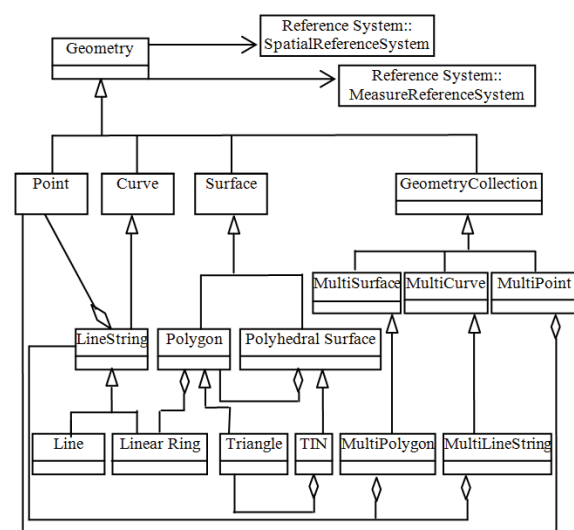


Fig. 3 - Hierarquia das Classes de Geometria.

Por exemplo, o seguinte comando:

```
SELECT geometria FROM município WHERE  
ST_DWithin (ST_GeometryFromText ('  
'POINT(-51.5 -21)', 4674));geometria, 100000);
```

retorna os polígonos dos municípios que estão a até 100 km de um ponto de coordenadas geográficas: longitude = -51° 30' e latitude = -21°, no sistema de referência SIRGAS 2000.

2.3 Buscas espaciais através de serviços web

Além de interagir diretamente com a informação geoespacial armazenada em Banco de Dados Geográficos, através de software específico, a disseminação do uso da Internet possibilitou que essa interação acontecesse também via web. Para tanto, se aplica a Arquitetura Orientada a Serviços (Service-Oriented Architecture – SOA), um estilo de arquitetura de sistemas cujo objetivo é a junção entre agentes de software que interagem entre si (HE, 2003). Para que isso aconteça, as diversas funcionalidades são disponibilizadas através de serviços interoperáveis.

Para garantir o bom funcionamento desta integração, uma série de normas foi criada pela OGC e ISO/TC211 (Comitê Técnico especializado em informação geográfica da Organização Internacional para Padronização – ISO), com base em padrões mais gerais desenvolvidos pela W3C, Consórcio World Wide Web - consórcio internacional de desenvolvimento de padrões para a Web. No caso da busca por dados geográficos na Internet, detalharemos três normas principais, a ISO 19136 - Geography Markup Language (GML), a ISO 19142 - Web Feature Service (WFS) e a ISO 19143 – Filter Encoding (FES).

A norma Web Feature Service (WFS) representa uma mudança na forma que a informação geográfica é criada, modificada e compartilhada na Internet (OGC, 2010c). No lugar da simples troca de arquivos, o serviço oferece acesso à informação ao nível de feições e suas propriedades. Os dados podem ser recuperados ou modificados usando um protocolo HTTP, no formato GML (Geography

Markup Language). Este formato deriva do XML (Extendible Markup Language), uma linguagem de marcação recomendada pelo W3C. É uma linguagem de propósito geral usada para criar outras para fins específicos, como o próprio GML, para informações geográficas, cujos conceitos principais são provenientes da série de normas ISO 19100 e da OpenGIS Abstract Specification.

As principais operações disponíveis para o serviço WFS são:

- GetCapabilities – é um chamado para determinar quais opções estão disponíveis em determinado serviço.
- GetFeature – é a operação que executa a obtenção dos dados em si. Parâmetros como o retângulo envolvente e outros devem ser informados e o serviço retornará o arquivo dos dados em formato GML com as geometrias e atributos pertinentes.

Dentro da operação GetFeature é possível restringir o resultado dos dados recuperados através da codificação de filtro (Filter Encoding, ou FES) definida pela norma ISO 19143 (OGC, 2010a). Este é o padrão definidor da sintaxe das expressões, descrevendo uma linguagem que proporciona a filtragem de dados baseados em seus identificadores, propriedades específicas e propriedades geoespaciais. Permite também a combinação de filtros usando expressões lógicas e a codificação de expressões aritméticas, dentre outras funções.

O Filter Encoding é também definido em XML, e foi projetado para ser originalmente parte da norma WFS, mas posteriormente foi definido em um documento a parte, para que pudesse ser usado como filtro em outros serviços, como o Web Coverage Service, Gazetteer e Web Registries (OGC, 2010a).

Um filtro é usado para identificar um subconjunto de recursos de uma coleção cujos valores das propriedades satisfaçam uma série de predicados conectados de forma lógica (OGC, 2010a). Se o valor da propriedade satisfizer todos os predicados do filtro, então esse recurso é considerado parte do subconjunto resultado. Os operadores que podem ser usados nos filtros são classificados em vários tipos conforme figura 04:

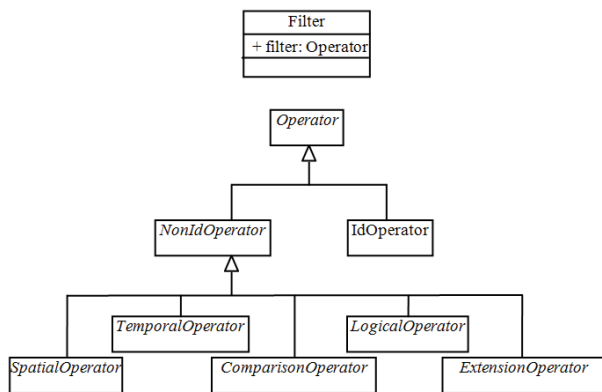


Fig 4 - Filtros.

Por exemplo, o comando descrito no item anterior, que selecionava os municípios com mais de 100 mil habitantes no Paraná, seria codificado assim:

```

<fes:And>
  <fes:PropertyIsEqualTo>
    <fes:PropertyName>Nome</fes:Property
Name>
    <fes:Literal>PR</fes:Literal>
  </fes:PropertyIsEqualTo>
  <fes:PropertyIsGreaterThan>
    <fes:PropertyName>Populacao
</fes:PropertyName>
    <fes:Literal>100000</fes:Literal>
  </fes:PropertyIsGreaterThan>
</fes:And>
  
```

O segundo exemplo, para o filtro espacial selecionando os municípios a até 100 km de um determinado ponto, seria feito como abaixo:

```

<fes:DWithin>
  <fes:ValueReference>geometry</fes:ValueReference>
  <gml:Point gml:id="P1" srsName="urn:ogc:def:crs:EPSG::4674">
    <gml:pos>-51.5 -21</gml:pos>
  </gml:Point>
  <fes:Distance uom="km">100</fes:Distance>
</fes:DWithin>
  
```

As buscas feitas usando o padrão FES podem ser facilmente transformadas para consultas SQL e vice-versa (GEOCONNECTIONS, 2012), bastando uma reestruturação entre os termos.

2.4 Busca usando WFS e/ou SQL

Uma ferramenta de Busca Tradicional (para diferenciar da Busca Semântica descrita no próximo item) segue a arquitetura conforme figura 05:

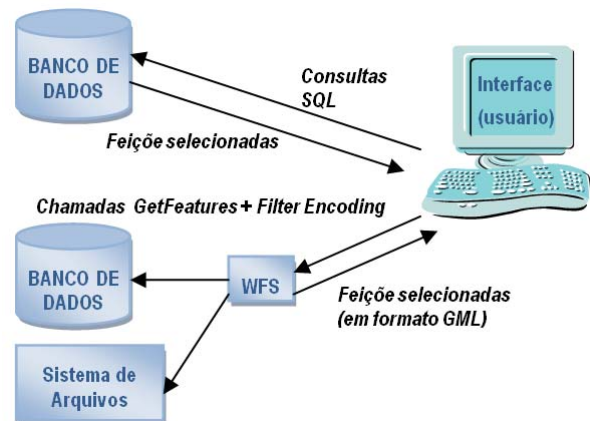


Fig. 5 - Arquitetura para uma ferramenta de Busca.

A seguir, são ilustrados quais os passos que um usuário deve seguir para realizar uma tarefa de busca de informações usando os dados disponíveis hoje na INDE-BR e no site do IBGE (www.ibge.gov.br). O exemplo foi escolhido por ilustrar tanto a Divisão Territorial Brasileira, informação de referência para diversas aplicações, e dados do Censo Demográfico, a única pesquisa a visitar todos os domicílios do Brasil e que fornece informações bastante ricas e detalhadas.

No exemplo, um usuário gostaria de saber a população de uma localidade chamada “Brejo de São Félix” a partir de um mapa de localização (perímetro). Ele não sabe se é uma sede municipal, distrital ou apenas um lugarejo isolado. Usando mecanismos de busca tradicional, ele precisaria seguir os seguintes passos.

1 – Acessar o serviço WMS “Localidades” disponibilizado pelo IBGE na INDE, de geometria “pontos”, provenientes da Base Cartográfica ao Milionésimo.

2 – Com a ferramenta de identificação no visualizador WMS utilizado, clicar sobre um ponto para descobrir que os nomes estão armazenados em um campo chamado “[nomeloc]”.

3 - Gerar a consulta SQL

```
SELECT * FROM localidades WHERE nomeloc = 'Brejo de São Félix'
```

4 – Como resultado, obtém-se a seguinte informação:

Brejo de São Félix-MA - localidades
gid: '1336'
nomemun: 'Parnarama'
tipo: 'Vila'
nomeloc: 'Brejo de São Félix'
anocria: '1948.00000'
estado: 'MA'

5 – Portanto, Brejo de São Félix é uma vila do município de Parnarama – MA.

6 – Para o próximo passo, o usuário precisaria obter a informação de que o termo “Vila” se refere à denominação oficial para uma Sede Distrital.

7 – Como a malha distrital não está publicada em forma de Geoserviço, o usuário precisa obter o shapefile da malha distrital do Estado do Maranhão (ftp://geoftp.ibge.gov.br/malhas_digitais/censo_2010/setores_censitarios/shape/ma_v1.3.zip)

8 - Adicionalmente, é preciso obter a informação que os arquivos correspondem às seguintes camadas:

21SEE250GC_SIR.shp = Setores Censitários
21SDE250GC_SIR.shp = Subdistritos
21MUE250GC_SIR.shp = Municípios
21DSE250GC_SIR.shp = Distritos

9 – Portanto, o usuário deve fazer uma nova consulta SQL no shapefile 21SDE250GC_SIR:

```
SELECT * FROM distritos WHERE NM_DISTRIT = 'BREJO DE SÃO FÉLIX'
```

Com o resultado, tem-se que este distrito possui o código CD_GEOCODI = '210780310'.

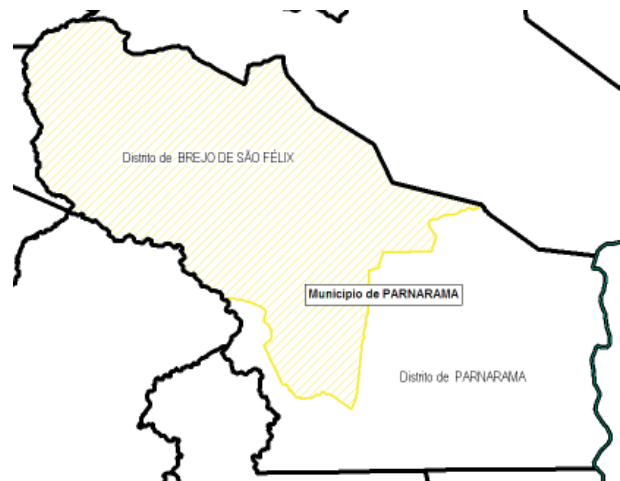


Fig. 6 - Distrito de Brejo de São Félix.

10 – Agora é possível selecionar os setores censitários que formam o distrito de Brejo de São Félix, com a consulta SQL no shapefile 21SEE250GC_SIR:

```
SELECT * FROM setores WHERE CD_GEOCODI like '210780310%'
```

Que obtém o resultado

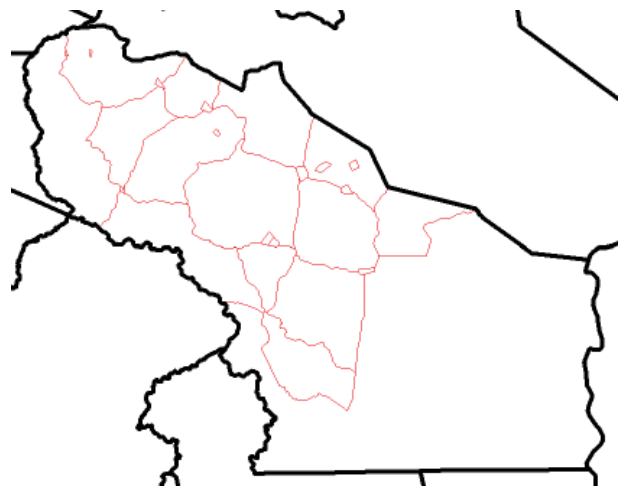


Fig. 7 - Setores Censitários do Distrito de Brejo de São Félix.

11 – Destes setores censitários, uma nova consulta para obter a sede do distrito (é necessário mais uma informação adicional: que a sede do distrito é representada pelo setor 01):

```
SELECT * FROM setores WHERE CD_GEOCODI = 210780310000001)
```

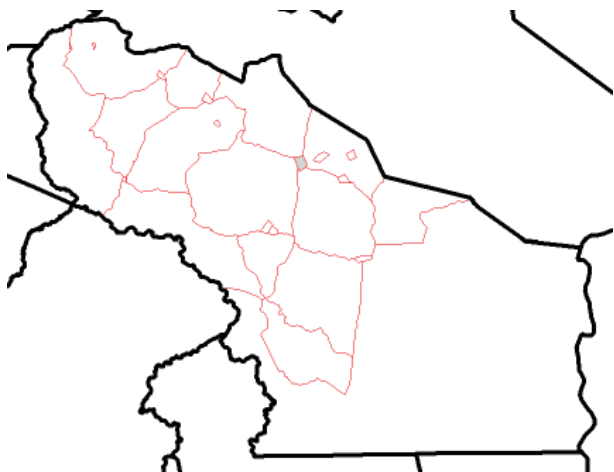


Fig. 8 - Sede do Distrito de Brejo de São Félix.

12 – Agora, com o código do Setor Censitário, é possível baixar as tabelas do resultado do Censo 2010 (ftp://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Sinopse/Agregados_por_Setores_Censitarios/Base_informacoes_setores2010_sinopse_MA.zip) e chega-se à informação desejada, ou seja, que a vila de Brejo de São Félix possuía, no Censo 2010, 715 residentes.



Fig. 9 - Brejo de São Félix – MA (715 habitantes – Censo 2010), sobre serviço de imagens Google Earth.

São 12 passos para a obtenção da informação, exigindo conhecimento da estrutura e

nomenclaturas de arquivos e tabelas, dos conceitos e codificação empregados e da localização de informação distribuída em diferentes páginas da web.

3. BUSCA GEOESPACIAL SEMÂNTICA

3.1 Heterogeneidade Semântica

Apesar do rápido desenvolvimento dos padrões e das tecnologias de serviços web ter sem dúvida ajudado o compartilhamento de informação geoespacial entre diversas fontes, a interoperabilidade resultante não pode resolver os problemas de heterogeneidade semântica (ZHAO et al, 2008).

Como indicado por CAI et al, 2007 o conhecimento espacial humano é incompleto, tendencioso, vago e algumas vezes inconsistente. Graças a esta condição, tanto a classificação que gerou a modelagem quando os critérios da busca são muito dependentes do contexto e da escala de representação. O conceito de proximidade, por exemplo, pode ser dependente do meio de transporte. Em uma busca “Quais são os restaurantes próximos ao teatro?” os resultados próximos são dependentes da forma de execução do percurso, se a pé ou de carro, e em que horário, por exemplo.

O desafio da estruturação semântica é conseguir retratar a visão de mundo empregada na modelagem dos dados para que este conhecimento possa ser aproveitado quando da utilização da informação. Nesta abordagem está sendo proposta a chamada Web Semântica, como mostrado na visão de BENNERS-LEE et al., 2001, que prevê a inclusão de estruturas semânticas na web. Com uma ligação estreita com o universo ontológico, desenvolveu-se a Engenharia do Conhecimento (Knowledge engineering), definida já em 1983 por FEIGENBAUM e McCORDUCK como: “a disciplina que envolve a integração de conhecimento em sistemas computacionais para resolver problemas complexos que normalmente requerem um alto nível de perícia humana”.

Dada a complexidade de se retratar o espaço geográfico computacionalmente, foi natural a expansão das pesquisas no campo da semântica para os dados espaciais. Dentre os trabalhos pioneiros que destacam problemas semânticos na informação geográfica está o

de WORBOYS e DEEN 1991, que já previa as vantagens de dados distribuídos em “nós” locais (antevendo a utilização da web) e define heterogeneidade semântica como o “desacordo a respeito do significado, interpretação ou uso pretendido para o dado”, dividida em dois grupos:

- Genérica: dados em modelo vetorial ou matricial, por exemplo – que vem se mostrando menos importante depois dos diversos esforços do setor em aspectos de interoperabilidade de dados e serviços geográficos;
- Contextual: causada por conceituações diferentes utilizadas em cada nó – que é mais similar ao significado atual do termo.

WACHE et al, 2001, percebe três divisões principais nos conflitos semânticos:

- Conflitos de confusão: ocorrem quando dois itens parecem ter o mesmo significado, mas na realidade são diferentes;
- Conflitos de escala: ocorrem quando diferentes sistemas de referência são usados para a medição de um valor, como por exemplo, diferentes escalas geográficas;
- Conflitos de nome: ocorrem quando nomes em esquemas de informação diferem significativamente, como por exemplo, a presença de sinônimos.

Outro artigo fundamental é o publicado por BENNETT em 2001: “What is a Forest”, que estudou as dificuldades de integração de conceitos geográficos, chamando atenção para a falta de destaque sobre questões sobre incerteza e imprecisão das informações no interesse dos geógrafos sobre os temas ontológicos. O OGC (Open Geospatial Consortium) já conduziu experimentos sobre o assunto: o “Geosemantics Interoperability Experiment” (REED e REICHARDT, 2008), que conclui que um fator crítico para o sucesso de longo prazo das Infraestruturas de Dados Espaciais é a interoperabilidade semântica.

Desde o início da utilização do termo no âmbito da Inteligência Artificial, a semântica se tornou também um tópico importante de pesquisa na comunidade da Ciência da Informação Geográfica, em aplicações como a do processamento da informação (recuperação

ou integração) feito com base em significado no lugar de “meras palavras-chave” (FONSECA e SHETH, 2002). CÂMARA, 2005 também destaca a dificuldade do compartilhamento de informações semânticas nos Sistemas de Informação Geográficas atuais, com a falta da inclusão de um conjunto de conceitos com um conjunto de relações semânticas, o que ele chama de Geo-Ontologias.

3.2 Ontologias

O termo Ontologia começou a ser empregado por estudiosos como GUARINO, 1998, que iniciaram a utilização do termo no âmbito da inteligência artificial, definindo-o como um “vocabulário específico usado para descrever certa realidade, somado com um conjunto de premissas explícitas sobre o sentido pretendido pelas palavras no vocabulário”, portanto, apresentando uma concordância de conceitos entre as partes envolvidas. Como o modelo cognitivo do indivíduo é produto da sua cultura, linguagem, educação e experiência de vida (STOCK, 2008), então mesmo dentro da mesma comunidade de informação, as visões do mundo diferem, portanto é necessária uma formalização dos conceitos empregados quando do processo de representação.

A estruturação das Ontologias em três camadas como na figura 10 é proposta por KLEIN e PROBST, 2005. O nível superior abriga os conceitos mais gerais e abrangentes, estando mais próximos dos campos de estudo da filosofia e da cognição. Já o nível de domínio congrega especialistas em determinada área, que podem detalhar a utilização de determinados conceitos neste campo do conhecimento. O nível de aplicação é uma ontologia de domínio especificada para uma aplicação em particular, explicitada para ser a base de um serviço específico e fortemente correlacionada com o esquema conceitual da mesma. Assim, é oferecida uma consistência lógica que possibilita a integração de Ontologias diversas criadas para domínios e aplicações específicas, mas sem perder a capacidade de tradução e interoperabilidade entre elas.

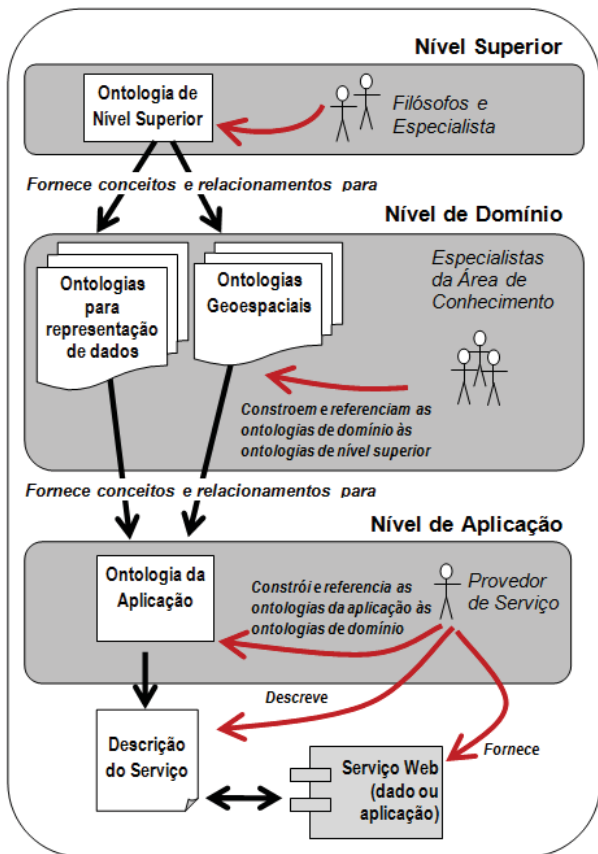


Fig. 10 - Arquitetura de Ontologias em 03 camadas.

Dada a complexidade de se retratar o espaço geográfico computacionalmente, foi natural a expansão das pesquisas no campo da semântica para os dados espaciais. As redes semânticas para informações geográficas incluem também relações topológicas como na Figura 11:

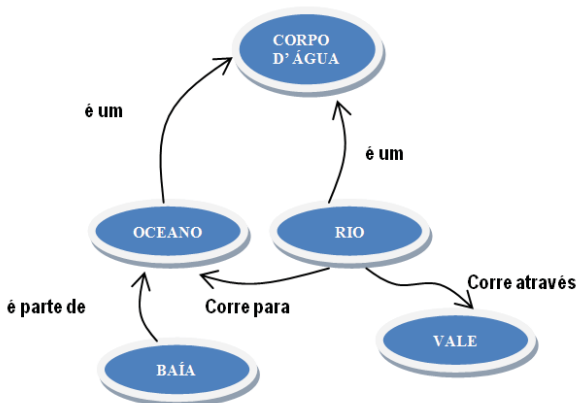


Fig. 11 - Rede Semântica Espacial

Desta forma uma ontologia que descreve relações espaciais permite buscas contendo também relações de vizinhança, pertencimento, conectividade e outras que aumentam o potencial de extração de conhecimento do modelo.

O W3C desenvolveu o padrão OWL (*Web Based Ontology Language*), uma linguagem para definir e instanciar ontologias na Web (SMITH et al., 2004). No final de 2009 esta especificação foi aprimorada e foi lançado o OWL 2, uma expansão e revisão dos conceitos da especificação anterior, de 2004. Outra subdivisão é o OWL-S, extensão da linguagem OWL para serviços web.

Uma ontologia OWL pode incluir descrições de classes, propriedades e suas instâncias, e é uma extensão de vocabulário do RDF (*Resource Description Framework*). O RDF é um padrão para intercâmbio de dados na Web também mantido pelo W3C. Ele se estrutura em declarações no formato sujeito-predicado-objeto ou recurso-propriedade-valor (conhecidos como triplas) que permite o compartilhamento de dados em diversas aplicações através da Web.

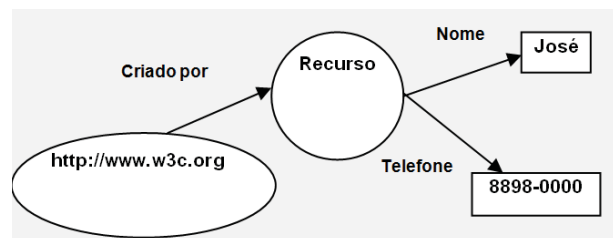


Fig. 12 - Exemplo de Triplas.

No exemplo da Figura 12 podem ser identificados:

- Classes: Página Web/ Autor
- Propriedades: Criado por/ Nome/ Telefone
- Valores: José/ 8898-0000

Enfim, é possível criar classes e subclasses, propriedades e subpropriedades, e associá-las entre si, como pode ser observado na Figura 13. As classes se referem a conceitos gerais e os indivíduos ou instâncias são os membros destes conjuntos. Muitas propriedades de conjuntos, como união e intersecção podem ser aplicadas nestes modelos.

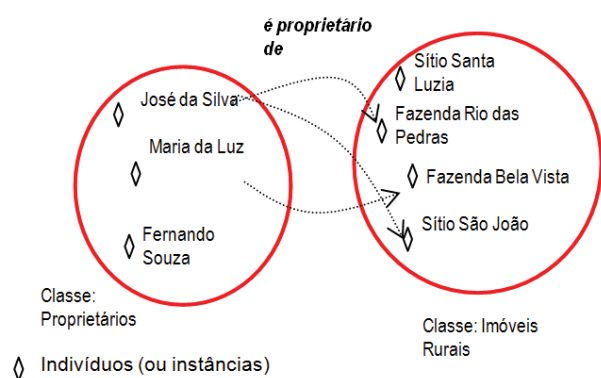


Fig. 13 - Classes e Indivíduos.

A criação dessas classes e a associação entre elas é feita com esforço conjunto das comunidades de produtores e usuários dos dados, bem como através de indexação de taxonomias e dicionários de dados já disponíveis.

3.3 Busca Semântica

Para se criar uma busca semântica em um Banco de Dados, a forma proposta, conforme também apresentada em diversas soluções (WEI et. al, 2008; SUN et al, 2009, ZHAO Et tal, 2008) é utilizar uma aplicação intermediária que, com base na solicitação original do usuário, use as ontologias para criar buscas SQL que possam ser endereçadas ao SGBD (ou via WFS, se através da web), onde as informações são geridas e armazenadas. A definição de ontologias neste caso possibilita a execução de tal tarefa, e é uma etapa fundamental no processo de construção da ferramenta.

A estrutura proposta está representada pela figura 14. Para possibilitar o uso das capacidades de modelos semânticos, são empregadas camadas de tradução que usam as ontologias para transformar as consultas do usuário em requisições reconhecíveis pelo banco de dados, ou requisições WFS que possam ser enviadas para serviços Web que então se comunicam através do padrão da OGC com suas próprias fontes de dados. WEI, 2007 adverte que apesar das buscas semânticas e SQL serem aparentemente próximas em termos de estrutura, a tradução sistemática e automática entre as duas é desafiadora, o que deve ser observado na implementação e testes da funcionalidade.

Observando a necessidade da definição da Ontologia para a execução da Busca Semântica, foi estabelecido abaixo a que seria necessária

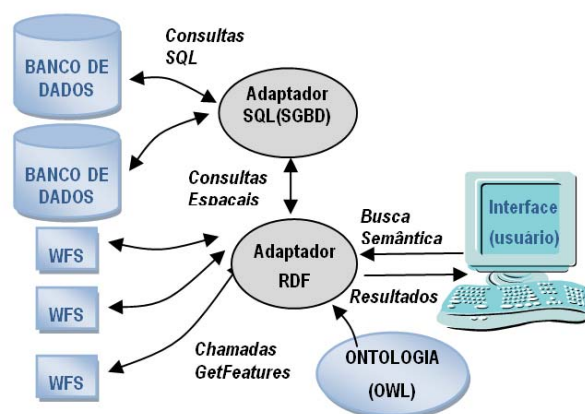


Fig. 14 - Arquitetura proposta para a ferramenta de Busca Semântica.

para resolução do exemplo descrito nos itens anteriores.

A descrição da Divisão Territorial Brasileira utilizada foi baseada em LIMA et al, 2002, buscando-se todas as inter-relações implícitas nas diversas hierarquias brasileiras para que a ontologia ficasse o mais completa possível. Existe uma diversidade grande de conceitos envolvidos, como por exemplo, determinada cidade, sede de um município, faz parte também de um distrito e subdistrito, é subdividida em bairros e este também em setores censitários, este definido como “*unidade espacial de coleta de informações utilizada pelo IBGE para a realização das pesquisas*” (LIMA et al, 2002), ao qual estão ligadas todas as tabelas de divulgação dos dados do Censo desagregados por setor.

De acordo com os dados do Censo 2010 do IBGE (IBGE, 2011), na data de referência do Censo existiam no Brasil 27 Unidades da Federação (sendo 26 Estados e 1 Distrito Federal), 5.565 municípios, 10.283 distritos, 662 subdistritos, 14.402 bairros e 316.574 Setores Censitários, os quais: 240.382 urbanos e 76.192 rurais. Essa diversidade de níveis administrativos e territoriais precisou ter seus relacionamentos mapeados de forma a criar a ontologia em formato OWL, usando o software livre Protégé, versão 4.2.0, desenvolvido pela Universidade de Stanford. Na figura 15 está o resultado em forma de gráfico.

Com esses relacionamentos disponíveis, o adaptador RDF seria responsável por disparar o *Reasoner* para que a consulta mostrada no item 2.4 pudesse ser feita de uma forma mais ágil. Procurando por “Brejo de São Félix” em

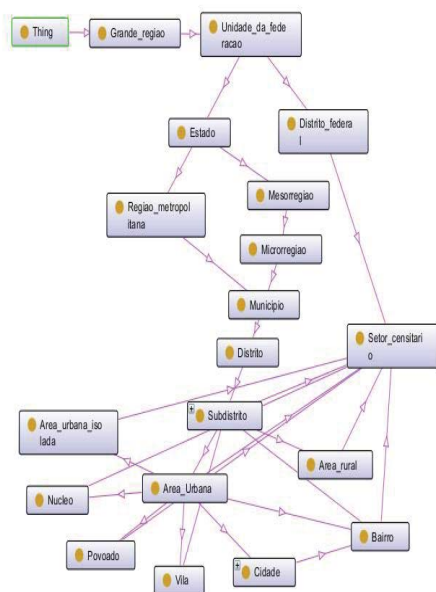


Fig. 15 - Ontologia para Divisão Territorial Brasileira.

todas as camadas disponíveis, automaticamente a identificaria como vila. A ligação ontológica com o setor censitário e tabela de população correspondente possibilitaria que o adaptador SQL pudesse criar diretamente as buscas nos bancos de dados, e trazer os resultados gráficos através de serviço WFS, por exemplo.

4. CONCLUSÃO

Em ambas as fases de busca por informações em uma IDE - descoberta dos dados através de metadados e recuperação das informações geoespaciais em si, a adição de capacidades semânticas aos mecanismos de busca pode auxiliar o usuário a encontrar a informação adequada, a interpretar os resultados e a identificar comportamentos e padrões nos mesmos.

Tanto o estado da arte da pesquisa na área quanto a análise de procedimentos de busca de dados apontam para os ganhos em usabilidade para as IDEs com a adição de recursos semânticos. A perspectiva é que a INDE-BR cresça em termos de dados e serviços disponíveis com a execução do Plano de Ação. Esta melhora em conteúdo não pode implicar em uma experiência pior para o usuário, mas ao contrário, pode fornecer as ferramentas necessárias para proporcionar um ambiente ágil e colaborativo.

As dificuldades principais começam por conseguir reunir diversos segmentos de produtores e usuários de dados em torno de projetos de criação de ontologias. Adicionalmente, é importante que sejam criadas aplicações que facilitem o processo de criação e troca de ontologias, incluindo as que procuram e indexam dados de outros serviços como LinkedIn, Dbpedia e outros, para que estas informações enriqueçam o conteúdo semântico da INDE-BR. Outra gama de aplicativos necessários são buscadores e aplicações de interfaces que permitem a utilização efetiva das ontologias estabelecidas.

Como as ontologias devem ser definidas desde o nível de domínio até o de aplicação, envolvendo produtores e usuários de múltiplas áreas, as ferramentas criadas devem ser testadas com usuários de perfis variados para mensurar os ganhos e obter subsídios para melhora dos processos. A combinação da necessidade de camadas de aplicação extra e do trabalho para definição das ontologias faz com que a Busca Semântica, tenha uma carga inicial de esforço de desenvolvimento maior, o que seria compensado pelos benefícios de ganho de usabilidade geral.

AGRADECIMENTOS

Nossos agradecimentos à equipe da Coordenação de Estruturas Territoriais (CETE), da Diretoria de Geociências do Instituto Brasileiro de Geografia e Estatística, em especial à equipe da Divisão Territorial Brasileira, pelo conhecimento compartilhado.

REFERÊNCIAS BIBLIOGRÁFICAS

- BENNETT, B., 2001. **What is a forest? On the vagueness of certain geographic concepts.** *Topoi*, 20(2), p.189–201.
- BENNERS-LEE, T., HENDLER, J. & LASSILA, O., **The semantic web.** Em: *Scientific American*, 284(5), p 34–43, 2001.
- BERNARD, L. & CRAGLIA, M., **SDI-from spatial data infrastructure to service driven infrastructure.** Em: *Research Workshop on Cross-Learning Between Spatial Data Infrastructures and Information Infrastructures*, Enschede, The Netherlands., 2005.
- CAI, G., **Contextualization of Geospatial Database Semantics for Human-GIS Interaction.** In: *Geoinformatica*, 11(2), 217-

237, 2007.

CÂMARA, G., **Representação Computacional de dados geográficos**. Em: Bancos de Dados Geográficos. Ed. Casanova, M.; Câmara, G.; Davis, C.; Vinhas, L.; Ribeiro de Queiroz, G.. Editora Mundogeo, Curitiba, p 1-44, 2005.

CONCAR. **Plano de Ação para Implantação da INDE**, 2010 Disponível em: < www.concar.ibge.gov.br/arquivo/PlanoDeAcaoINDE.pdf> [acessado em outubro, 2012]

CROMPVOETS, J. & BREGT, A., 2003. **World status of national spatial data clearinghouses**. Em: Journal of the Urban and Regional Information Systems Association, Volume 15• APA I•, p.43.

FEIGENBAUM, E. & MCCORDUCK, P. **The fifth generation : artificial intelligence and Japan's computer challenge to the world**, Addison-Wesley Pub. Co., Reading, MA, 1983, p.275

FERNÁNDEZ, T.D. & CASTELLANOS, E., **Towards user-driven spatial data infrastructures. An approach oriented to sustainable development**. Em: GSDI-9 World Conference, Santiago, Chile, 2006.

FERREIRA, K. R., CASANOVA, M. A., DE QUEIROZ, G. R. e DE OLIVEIRA, O. F., **Arquiteturas e linguagens**. Em: Bancos de Dados Geográficos . Ed. Casanova, M.; Câmara, G.; Davis, C.; Vinhas, L.; Ribeiro de Queiroz, G. Editora Mundogeo, Curitiba, 2005 pp. 181-212

FONSECA, F. & SHETH, A., 2002 **Research Priorities: The Geospatial Semantic Web** – The University Consortium for Geographic Information Science, Disponível em: http://www.ucgis.org/priorities/research/2002researchPDF/shortterm/e_geosemantic_web.pdf [acessado em outubro, 2012]

GRUS, L., CROMPVOETS, J. & BREGTA., 2006. **Defining National Spatial Data Infrastructures as Complex Adaptive Systems** Em: GSDI-9 World Conference Proceedings, Santiago, Chile, 2006. pp. 6-10

GUARINO, N. **Formal Ontology in Information Systems**. Em: Proceedings of FOIS'98, Trento, Italy. Amsterdam, IOS Press, 1998 pp. 3-15

IBGE, 2011 **Malha Territorial Censo 2010** Disponível em: <<http://www.ibge.gov.br/home/presidencia/noticias/imprensa/ppts/0000000483.pdf>> [acessado em outubro, 2012]

KLIEN, E. **Semantic Annotation of Geographic Information**. PhD Thesis, Institute for Geoinformatics, University of Muenster. Muenster, Germany, 2008

KLIEN, E., e PROBST, F. **Requirements for Geospatial Ontology Engineering**. Em: 8th Conference on Geographic Information Science (AGILE 2005). Estoril, Portugal, 2005. pp. 251-260

LEITE JR, F.L. et al., **WS-GIS: Towards a SOA-Based SDI Federation**. Em: VIII Brazilian Symposium on GeoInformatics, Campos do Jordão, 2006.

LIMA, M.H.P.(org); RODRIGUES, C.M.; SILVA, J.K.T.; MARTINS, P.C.; TERRON, S.L. e SILVA, R.L.deS., **Divisão Territorial Brasileira**. Disponível em: <http://www.academia.edu/336067/Divisao_Territorial_Brasileira> [acessado em outubro, 2012]

MASSER, I., **Changing Notions of a Spatial Data Infrastructure**. Em: SDI Convergence. Rotterdam, The Netherlands, p. 219, 2009.

OGC, **OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture (OGC 06-103r3)**, Version 1.2.0, 2006

OGC, **Filter Encoding Implementation Specification (OGC 09-026r1 e ISO 19143:2010)**, Version 2.0.0, 2010

OGC, **OpenGIS Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option (OGC 06-104r4)**, Version 1.2.1, 2010

OGC, **OpenGIS Web Feature Service 2.0 Interface Standard (OGC 09-025r1 e ISO/DIS 19142)**, Version 2.0.0, 2010

PAN, T.; ZHENG, L., ZHANG, H.; FANG, C.; LOU, J. & SHAO, Z. **Combining Web Services Toward Innovative Design of Agile Virtual Enterprise Supported by Web 3.0**. Em: WSEAS Transactions on Communications, Volume 8 Issue 1, 2009 pp 81-91

- REED, C. & REICHARDT, M.. **From Spatial Data Infrastructure to the Geospatial Web: Advancing the SDI Vision in a Changing Technology Environment.** Em: GSDI-9 World Conference Proceedings, St. Augustine, Trinidad , 2008. p. 30.
- SADEGHI-NIARAKI, A.; RAJABIFARD, A; KIM, K e SEO J. **Ontology Based SDI to Facilitate Spatially Enabled Society** In: Global Spatial Data Infrastructure 12th International Conference Proceedings, Singapore, 2010. pp. 19-22
- SILBERSCHATZ, A. ; KORTH, H. & SUDARSHAN, S. **Database System Concepts.** McGraw-Hill Science/Engineering/Math, 5 edition, 2005. 1168 p.
- SMITH, M.K., WELTY, C. & MCGUINNESS, D.L. **OWL Web Ontology Language Guide,** 2004 Disponível em: <<http://www.w3.org/TR/owl-guide/>> [acessado em outubro, 2012]
- STANFORD UNIVERSITY SCHOOL OF MEDICINE, **Protégé,** versão 4.2.0 Disponível em: <<http://protege.stanford.edu>> [acessado em outubro, 2012]
- STOCK, K., **Ontology-Schmology: In search of informal semantics.** Em: Workshop on Geographic Ontologies Grounding and Vagueness, Ilkley, West Yorkshire, UK, 2008. p. 4.
- SUN, C.H. et al., **Development of an Ontology-based Service Oriented Architecture Framework for National Geographic Information System in Taiwan.** In: Global Spatial Data Infrastructure 11th International Conference Proceedings, The Netherlands, 2009.
- WACHE , H.; VÖGELE, T. ; VISSER, U. ; STUCKENSCHMIDT, H.; SCHUSTER, G.; NEUMANN, H. & HÜBNER, S. **Ontology-based integration of information-a survey of existing approaches.** Em IJCAI-01 Workshop: Ontologies and Information Sharing. pp. 108–117, 2001.
- WEI, M., **Enabling Semantic Power of Geographic Data** In: Ontology for The National Map Workshop, Feb. 3-4, 2009. Disponível em: < <http://cegis.usgs.gov/pdf/Wei2009EnablingSemanticPower.pdf>> [acessado em outubro, 2012]
- WEI, M., ZHAO, T., VARANKA, D., e USERY, L. **A conceptual design towards semantic geospatial data access.** Em: 5th International Conference on Geographic Information Science, Park City, Utah, 2008. p. 6
- WORBOYS, M.F. & DEEN, S.M. **Semantic heterogeneity in distributed geographic databases.** Em: ACM Sigmod Record, 20(4), pp 30-34, 1991.
- ZHAO, P., DILL., YU, G., YUE,P., WEI, Y., e YANG, W., **Semantic Web-based geospatial knowledge transformation.** Em: Computers & Geosciences 35, n. 4 pp 798-808, 2009.