

Infraestrutura Nacional de Dados Espaciais – Aspectos Computacionais

Jaudete Daltio¹

Claudio A. Spadotto²

Beatriz B. O. Santos³

Natália dos S. Fois⁴

Paulo R. R. Martinho⁵

Sâmara R. da S. Trajano⁶

Wilson A. Holler⁷

Introdução e motivação

Dados geográficos (ou geoespaciais) representam fatos, objetos e fenômenos associados a uma localização sobre a superfície terrestre, num certo instante ou período de tempo (CÂMARA et al., 1996a). Dados geográficos ganharam maior popularidade nas últimas décadas e hoje representam uma parcela significativa dos dados disponíveis na *Web*. Com o intuito de viabilizar a ampla utilização e o compartilhamento desses dados pelo Estado e pela sociedade, vários países iniciaram a definição, das chamadas Infraestrutura de Dados Espaciais (IDE) (BRASIL, 2010c). Dentre as principais motivações para a implementação de IDEs está a importância crescente de informações sobre a distribuição geográfica de diversos fenômenos no planejamento de atividades políticas

e administrativas para o desenvolvimento social, ambiental e econômico da sociedade. A inexistência de uma infraestrutura de dados espaciais integrada, aberta e atualizada pode também trazer retrabalho e outros impactos negativos ao andamento de projetos de diversas áreas de pesquisa (CASTRO, 2012).

No Brasil, o Decreto nº 6.666 de 27 de novembro de 2008 instituiu a IDE Nacional (INDE) e a definiu como o “conjunto integrado de tecnologias; políticas; mecanismos e procedimentos de coordenação e monitoramento; padrões e acordos, necessário para facilitar e ordenar a geração, o armazenamento, o acesso, o compartilhamento, a disseminação e o uso dos dados geoespaciais de origem federal, estadual, distrital e municipal” (BRASIL, 2008). Num país de dimensão continental como o Brasil, com uma grande carência de

¹ Analista de Sistemas, doutoranda em Ciência da Computação, analista da Embrapa Gestão Territorial, Campinas, SP. E-mail: jaudete.daltio@embrapa.br

² Engenheiro-agrônomo, Ph.D. em Ciência do Solo e Água, pesquisador da Embrapa Gestão Territorial, Campinas, SP. E-mail: claudio.spadotto@embrapa.br

³ Graduanda em Geografia, estagiária da Embrapa Gestão Territorial, Campinas, SP. E-mail: beatriz.santos@colaborador.embrapa.br

⁴ Analista de Sistemas, mestre em Engenharia Civil, analista da Embrapa Gestão Territorial, Campinas, SP. E-mail: natalia.fois@embrapa.br

⁵ Engenheiro-agrônomo, mestre em Agricultura Tropical, Subtropical e Gestão de Recursos Ambientais, analista da Embrapa Gestão Territorial, Campinas, SP. E-mail: paulo.martinho@embrapa.br

⁶ Geógrafa, mestre em Manejo de Solo e Água, analista da Embrapa Gestão Territorial, Campinas, SP. E-mail: samara.trajano@embrapa.br

⁷ Engenheiro-cartógrafo, especialista em Geoprocessamento, analista da Embrapa Gestão Territorial, Campinas, SP. E-mail: wilson.holler@embrapa.br

informações adequadas para a tomada de decisões sobre os problemas urbanos, rurais e ambientais, esse compartilhamento de dados geográficos apresenta um enorme potencial, aumentando a eficiência da gestão territorial estratégica e influenciando o planejamento, a implantação e o acompanhamento das políticas públicas e de setores privados (SPADOTTO, 2012).

A INDE, assim como toda IDE, é fundamentada em cinco principais pilares que são fortemente relacionados e interagem entre si (WARNEST, 2005). A Figura 1 apresenta essa organização, que tem os “dados” (*Data*) como seu pilar principal, representando conjuntos de dados geoespaciais. O pilar “pessoas” (*People*) representa as partes (setores público, privado ou acadêmico) envolvidas em todo o processo. O pilar “institucional” (*Institutional Frameworks*) abrange questões políticas, de legislação e coordenação que detalham os termos e as condições para o uso de dados geoespaciais. O pilar “tecnologia” (*Technology*) descreve as tecnologias de infraestrutura necessárias para o estabelecimento da IDE, bem como os mecanismos computacionais para viabilizar o acesso, o armazenamento e a consulta de dados geoespaciais. O último pilar, “padrões” (*Standards*), serve de base para o pilar de tecnologia e permite o intercâmbio e a integração desses dados. Esses padrões abrangem modelos e dicionários de dados, padrões de intercâmbio de dados e metadados.

O objetivo deste documento é apresentar os principais conceitos relacionados aos pilares

Fundamental Datasets	Data
Users, Providers, Communication, Collaboration	People
Policy, Legislation, Coordination	Institutional Frameworks
Access, Acquisition, Distribution, Storage	Technology
Data Models, Metadata, Transfer Standards	Standards

Figura 1. Pilares de uma IDE.
Fonte: Warnest (2005).

“tecnologia” e “padrões”, com enfoque em seus aspectos computacionais. A intenção é sumarizar os fundamentos teóricos envolvidos na especificação da INDE e reunir as principais referências de softwares *open source* atualmente disponíveis e amplamente adotados pela comunidade para a implantação da INDE.

O restante deste documento está organizado como segue. A seção 2 descreve os principais conceitos e padrões que servem de embasamento para as definições tecnológicas utilizadas na IDE, ressaltando os padrões de interoperabilidade técnica e semântica, e os padrões para integração de dados via serviços *Web*. Em sequência, a seção 3 aborda as principais soluções tecnológicas de software *open source* que implementam esses padrões de interoperabilidade e viabilizam a implantação de uma IDE baseada em software livre. A última seção descreve as considerações finais sobre o estado atual dessas tecnologias e os principais desafios a serem superados nos próximos anos.

Padrões da INDE – o e-PING

O principal motivador para o estabelecimento de padrões como uma das camadas da IDE é a possibilidade de prover interoperabilidade, ou seja, permitir a interação e o intercâmbio de dados de acordo com um método e um formato bem definidos. Interoperabilidade é a soma de vários fatores que consideram a integração de sistemas e de redes, a troca de dados entre sistemas e a definição de tecnologias partindo do pressuposto da diversidade de componentes e plataformas. Seu objetivo é considerar todos os fatores para que os sistemas possam atuar cooperativamente, fixando as normas, as políticas e os padrões necessários.

O e-PING (Padrões de Interoperabilidade de Governo Eletrônico) define um conjunto mínimo de premissas, políticas e especificações técnicas que regulamentam a utilização da tecnologia de informação na interoperabilidade de serviços de governo eletrônico (BRASIL, 2010b). Seu propósito é estabelecer políticas e especificações técnicas que permitam a prestação de serviços eletrônicos de qualidade à sociedade.

As áreas cobertas pelos padrões e-PING estão segmentadas em: (i) interconexão; (ii) segurança; (iii) meios de acesso; (iv) organização e intercâmbio de informações; e (v) áreas de integração

para governo eletrônico. Para cada um desses segmentos, foram especificados componentes e padrões fundamentados em políticas como, por exemplo, a adoção de padrões abertos nas especificações técnicas, o uso de software público e/ou livre na implementação dos padrões, a adoção de soluções apoiadas pelo mercado, o alinhamento com as especificações utilizadas na *Web*, a adoção do XML como padrão de intercâmbio de dados, a adoção de navegadores (*browsers*) como principal meio de acesso, o desenvolvimento e a adoção de um padrão de metadados e a organização e recuperação de informações por meio de vocabulários controlados, taxonomias e ontologias.

As seções subsequentes descrevem os principais conceitos e padrões computacionais abrangidos pelos segmentos (iv) e (v).

Organização e intercâmbio de informações

Esse segmento do e-PING aborda os aspectos relativos ao tratamento e à transferência de informações nos serviços de governo eletrônico e inclui aspectos técnicos e semânticos da representação desses dados e seus metadados.

Interoperabilidade técnica

A interoperabilidade técnica está relacionada à adoção de linguagens para definição, intercâmbio e transformação de dados. A linguagem base adotada no e-PING para atender a esse propósito é a *Extensible Markup Language* (XML) (BRAY et al., 2000). XML é uma metalinguagem de marcação recomendada pelo *World Wide Web Consortium* (W3C)⁸ que permite a definição de linguagens de marcação para domínios específicos. O XML é uma linguagem de dados textual que define uma série de regras para a representação de dados em um formato que seja legível para máquina e humanos, sendo amplamente utilizada para a troca de informações de forma estruturada pela *Web*.

A linguagem XML enfatiza a simplicidade, a generalidade e a usabilidade. Muitas interfaces de programação de aplicativos (*application programming interface* – APIs) foram desenvolvidas para usar e processar dados XML, e vários sistemas

se baseiam em XML para a definição de suas linguagens. Ao criar uma infraestrutura única para outras linguagens, o XML permite que novas linguagens possam ser definidas sem maior esforço e sem necessidade de submissão aos comitês de padronização.

Interoperabilidade semântica

A interoperabilidade semântica busca a definição de elementos que permitam a interpretação dos dados e o enriquecimento de seu contexto. A intenção é subsidiar o uso adequado dos dados por diferentes tipos de usuários propiciando a compreensão e o entendimento sobre a sua aplicabilidade e forma de utilização. Nesse sentido, duas vertentes principais são propostas pelo e-PING: a utilização e padronização de metadados e o uso de vocabulários controlados.

Metadados são comumente definidos como dados sobre os dados, informações estruturadas utilizadas para descrever um recurso ou parte dele. Metadados geoespaciais (MG) representam uma classe particular de metadados que permite caracterizar informações geoespaciais. Esses metadados são imprescindíveis para o funcionamento adequado da INDE, pois permitem que os usuários possam identificar potenciais conjuntos de dados de interesse.

No escopo da INDE, prevê-se a utilização do Padrão de Metadados para o Governo eletrônico (e-PMG) (BRASIL, 2010a) para a representação de metadados. O e-PMG foi baseado no padrão *Dublin Core* (DC) (WEIBEL; KOCH, 2000), uma organização engajada no desenvolvimento de um padrão de metadados interoperável que busca a simplicidade, a semântica universal, o escopo internacional e a extensibilidade. O e-PMG estabelece um conjunto mínimo de elementos que contém dados necessários para a recuperação e o gerenciamento de informações que consiste em 20 elementos, 15 deles do padrão DC, que são:

- Abrangência: extensão espacial e temporal do recurso.
- Assunto: palavras-chaves ou termos que representem corretamente o conteúdo do recurso.
- Colaborador: entidade que contribuiu para a elaboração do conteúdo do recurso ou para sua divulgação.

⁸ Disponível em: <<http://www.w3.org/>>.

- Criador: entidade originalmente responsável pela criação e manutenção do conteúdo do recurso.
 - Data: data associada a um evento no ciclo de vida do recurso.
 - Descrição: resumo e explicação do conteúdo do recurso.
 - Direitos: informação relativa às restrições e permissões sobre o acesso e o uso do recurso, tais como ler, copiar, distribuir, publicar ou ainda fazer uso de todo ou de parte do recurso.
 - Fonte: identificação do recurso que deu origem ao recurso atual.
 - Formato: conjunto de características físicas e lógicas que define a apresentação do recurso.
 - Identificador: referência inequívoca ao recurso em um determinado contexto.
 - Idioma: idioma em que é expresso o conteúdo do documento.
 - Publicador: organização ou pessoa responsável por tornar o recurso disponível, o qual pode ser uma edição impressa ou uma publicação eletrônica.
 - Relação: referência a um outro recurso relacionado.
 - Tipo: natureza ou gênero do conteúdo do recurso.
 - Título: palavra, frase ou grupo de caracteres que nomeia um recurso e que pode corresponder à sua denominação oficial ou legal, ou ainda ao nome que aparece de forma evidente ou explícita.
- arquivísticos para guarda permanente ou eliminação.
 - Destinatário: entidade para quem a informação contida no documento foi dirigida;
 - Localização: local de armazenamento atual do recurso.
 - Preservação: informação para apoiar o acesso, o uso e a preservação de um recurso a longo prazo.

Adicionalmente, foram estabelecidos outros 5 elementos identificados como necessários para o contexto do governo eletrônico brasileiro:

- Contexto jurídico-administrativo: legislação e normas que regulam e autorizam o funcionamento e o desempenho das atividades da organização responsável pela produção e/ou disponibilização do recurso.
- Destinação: decisão quanto ao encaminhamento dos documentos

O e-PMG utiliza o Vocabulário Controlado do Governo Eletrônico (VCGE) (BRASIL, 2011a) como base para criação de categorias para o elemento “assunto”, possibilitando a indexação de serviços e informações governamentais. O VCGE é um esquema de termos intuitivos para os cidadãos que buscam assuntos do seu interesse na larga faixa de informações do setor público e não requer o conhecimento do órgão responsável pelo assunto. A Figura 2 ilustra a taxonomia para navegação do VCGE, que possui, em seu primeiro nível, conceitos como “indústria”, “meio ambiente” e “trabalho”. Como pode ser visto, o tema “área de preservação permanente (APP)” é um termo mais específico oriundo de “áreas protegidas”, que é agrupado com vários outros termos (“águas”, “florestas” e “biodiversidade”, por exemplo) no conceito “meio ambiente”.

Embora a adoção de metadados e de um vocabulário controlado garanta muitos benefícios na descrição de informações, esses mecanismos ainda são insuficientes para prover a interoperabilidade semântica necessária na INDE. De forma a complementar essa semântica, o e-PING propõe os vocabulários e ontologias do governo eletrônico (e-VoG) (BRASIL, 2011c). Trata-se de um conjunto de padrões, ferramentas e metodologias com o objetivo de possibilitar o intercâmbio de informações com acordo semântico, viabilizando o cruzamento de dados de diversas fontes, o uso de metodologias de modelagem conceitual como forma de obtenção do conhecimento tácito das áreas de negócio de governo e o alinhamento conceitual das diversas áreas do conhecimento do governo.

Para alcançar essas metas, o e-VoG prevê o uso de ontologias para explicitar conhecimentos de maneira formal e coerente. Ontologias são

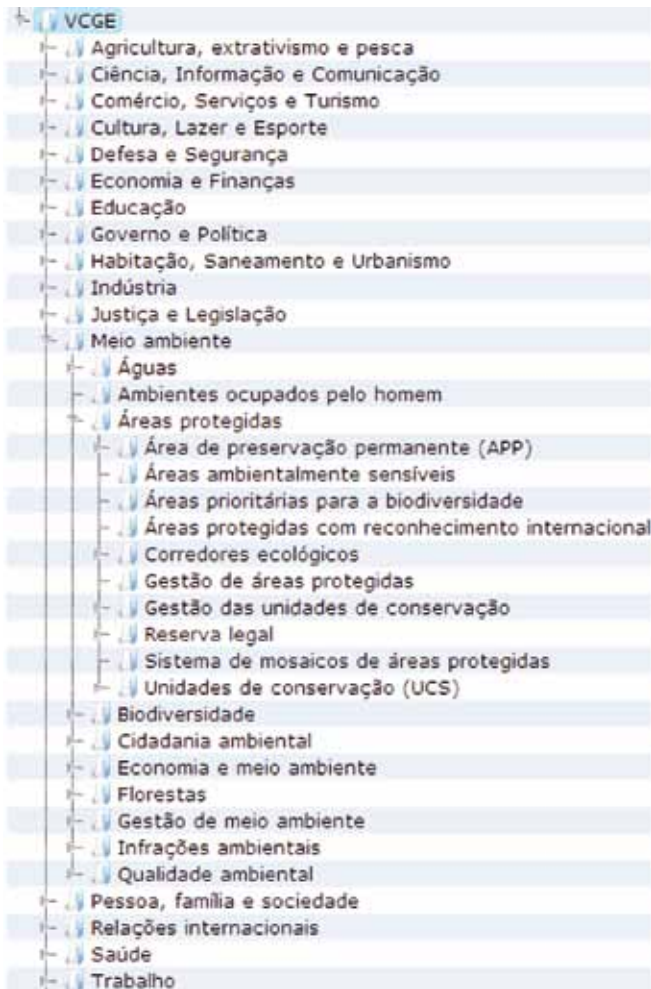


Figura 2. Taxonomia de categorias VCGE.

Fonte: Brasil (2011b).

estruturas computacionais capazes de representar um modelo abstrato de termos relacionados entre si (HORROCKS, 2008). Elas modelam uma parte da realidade, os termos de um domínio, suas entidades, relações taxonômicas e não taxonômicas e as restrições aplicadas a essas entidades de domínio, com a finalidade de definir um entendimento comum sobre um domínio. A partir do domínio estruturado disponibilizado pelas ontologias para representar as relações entre seus termos, torna-se possível interpretá-los de forma clara, livre de ambiguidades. O resultado é um entendimento comum e legível para processamento automático. O uso de ontologias permite diferenciar, por exemplo, se o termo “vírus” corresponde a um microrganismo, caso o contexto seja biológico ou agrícola, ou se corresponde a um programa de computador malicioso, caso o contexto seja computação.

As ontologias desenvolvidas no âmbito do e-VoG farão referência a conceitos externos definidos em

ontologias que tenham ampla utilização nacional e internacional, de forma a aumentar o potencial para o cruzamento de dados de outras fontes.

Áreas de integração para governo eletrônico

Esse segmento do e-PING estabelece a utilização de especificações técnicas baseadas no padrão XML para promover o intercâmbio de informações em áreas transversais da atuação governamental. O padrão de intercâmbio de dados proposto na INDE é fundamentado pela arquitetura orientada a serviços (sigla em inglês: SOA), utilizando os serviços *Web* como principal instância desse paradigma. As seções subsequentes descrevem esses conceitos, incluindo suas especificidades para a manipulação de dados geoespaciais.

Arquitetura Orientada a Serviços (SOA)

SOA é um estilo arquitetural cujo objetivo é alcançar fraco acoplamento entre softwares, disponibilizando-os como serviços que interagem entre si. Serviços são elementos computacionais que suportam composição rápida, e a baixo custo de aplicações distribuídas, e são independentes e autodescritivos (PAPAZOGLU, 2003). Esses serviços podem realizar tarefas que vão desde simples requisições de busca até complexos cálculos matemáticos.

Em SOA, as funcionalidades das aplicações são agregadas como serviços reutilizáveis, cada qual disponibilizado por uma interface que especifica as funcionalidades, as condições de exceção, os comandos aceitos e as respostas a tais comandos. Todas as requisições, respostas, condições de exceção e funcionalidades devem ser listadas como parte dessa interface. Os serviços são disponibilizados pelos fornecedores de serviços e utilizados pelos consumidores de serviços. As três principais características de uma arquitetura orientada a serviços são (PAPAZOGLU, 2003):

- Independência de plataforma: serviços devem ser acessáveis utilizando tecnologias existentes em ambientes distintos. Isso implica que protocolos e mecanismos de descrição e recuperação de serviços devem seguir padrões formalmente definidos e aceitos.

- **Acoplamento fraco:** serviços não devem conhecer a estrutura e o funcionamento interno de outros serviços. O acesso é sempre intermediado pela interface do serviço. Isso aumenta a flexibilidade da arquitetura, uma vez que a alteração interna de um serviço não afeta seus clientes.
- **Transparência de localização:** serviços devem poder ser encontrados pelos clientes independentemente de sua localização. Suas definições e informações de localização devem ser armazenadas em um diretório de serviços que é utilizado pelo cliente na busca.

Serviços Web

Serviços *Web* representam a tecnologia de implementação mais utilizada em SOA. São aplicações que podem ser encontrados, descritos e acessados via protocolos *Web* padrão. Se construídas como Serviços *Web*, aplicações escritas em diversas linguagens e rodando sobre diferentes plataformas estão habilitadas a interagir e trocar informações de forma transparente (ALONSO et al., 2004).

Um Serviço *Web* é uma noção abstrata que deve ser implementada por um agente concreto. O agente pode ser um módulo de software ou um componente de hardware que envia ou recebe mensagens, e o serviço é um conjunto abstrato de funcionalidades (interfaces) que são disponibilizadas. Serviços *Web* são identificados por um *Uniform Resource Identifier* (URI) ou um *Uniform Resource Locator* (URL)⁹, descritos e definidos por linguagens baseadas em XML, e utilizam o protocolo *Hypertext Transfer Protocol* (HTTP)¹⁰ para o transporte de requisições. Serviços *Web* podem ser acessados por outras aplicações usando, por exemplo, o protocolo *Simple Object Access Protocol* (SOAP) (BOX et al., 2000) por meio de suas interfaces públicas descritas, por exemplo, em *Web Services Description Language* (WSDL) (CHRISTENSEN et al., 2001).

1. Geo Serviços Web

O *Open Geospatial Consortium* (OGC)¹¹ é um consórcio internacional de empresas, agências governamentais e universidades que visa ao

⁹ Disponível em: <<http://www.w3.org/TR/uri-clarification>>.

¹⁰ Disponível em: <<http://www.w3.org/Protocols>>.

¹¹ Disponível em: <<http://www.opengeospatial.org/>>.

desenvolvimento de especificações de interfaces e padrões de intercâmbio de dados geoespaciais. Esses padrões visam à interoperabilidade e possibilitam a implementação de geo serviços *Web*, permitindo que serviços e informações espaciais complexas estejam acessíveis a todos os tipos de aplicações, independentemente da plataforma utilizada.

A ideia principal é oferecer serviços de acesso a dados geoespaciais com base em catálogos de dados, tornando indiferentes aos olhos do usuário o local, o meio e a estrutura física de armazenamento. O acesso aos dados é realizado apenas através de serviços, pelo encapsulamento da estrutura física dos dados. O usuário não precisa conhecer o local onde os dados estão armazenados, pois cada provedor de dados se encarrega de registrar, junto a um serviço de catalogação, que dados possui, onde estão, como estão organizados e onde estão os metadados. As estruturas para esses serviços devem ser distribuídas, suportar múltiplas aplicações, clientes de diversos tipos, fontes de dados variadas e múltiplos grupos para manutenção e atualização, compondo um ambiente computacional heterogêneo.

Os Serviços *Web* representam o componente central nessa arquitetura. Serviços individuais têm interfaces de tipos conhecidos, são descritas utilizando metadados de serviços e disponibilizadas para os usuários através de solicitações padronizadas pela OGC. Alguns dos serviços mais importantes especificados e documentados pela OGC (MICHAELIS; AMES, 2008) são descritos a seguir de forma resumida.

- **Web Map Service (WMS):** define a interface de um serviço para disponibilizar mapas que serão apenas uma representação visual dos dados espaciais (e não os dados em si). Essas representações são geradas em formatos de imagem como .jpeg, .png e .gif. O serviço WMS permite visualizar dados geográficos e consultar as entidades mostradas num mapa vetorial e sobrepor dados vetoriais e dados matriciais em diferentes formatos, sistemas de referência e sistemas de projeção, situados em diferentes servidores. As requisições WMS podem ser feitas por um navegador padrão em forma de URLs. As operações suportadas por esse serviço são:

- a) *GetCapabilities*: retorna os metadados do serviço descrevendo as interfaces suportadas pelo servidor, as camadas de informação disponíveis e os respectivos formatos possíveis.
- b) *GetMap*: retorna imagens que correspondem a uma determinada visualização estática, combinando uma ou mais camadas de informação.
- c) *GetFeatureInfo*: fornece mais informações a respeito de pontos específicos de uma imagem retornada pelo servidor.
- *Web Feature Service (WFS)*: define um serviço que permite acessar e manipular dados geográficos codificados em formato *geographic markup language (GML)* (SANTOS; LU, 2008). Permite ao usuário acessar, consultar e modificar (inserir, atualizar e eliminar) todos os atributos de um fenômeno geográfico representado em formato vetorial. As operações suportadas por esse serviço são:
 - a) *GetCapabilities*: retorna as interfaces suportadas pelo servidor e as camadas de informação disponíveis para requisições.
 - b) *DescribeFeatureType*: solicita as características e atributos de uma camada de informação.
 - c) *GetPropertyValue*: recupera o valor de uma propriedade de uma camada de informação utilizando uma expressão de consulta.
 - d) *GetFeature*: recupera os dados de uma camada de informação vetorial, incluindo sua geometria e os valores de seus atributos.
 - e) *LockFeature (WFS-T)*: impede que uma camada de informação seja editada. Essa operação é útil no processamento de transações, em que os usuários podem bloquear uma área do mapa para edição, garantindo que não haja outras edições simultâneas.
 - f) *GetFeatureWithLock (WFS-T)*: combinação das operações *GetFeature* e *LockFeature*, recupera e bloqueia os dados de uma camada de informação.
 - g) *Transaction (WFS-T)*: permite a edição dos dados de uma camada de informação existente por meio de inserções, atualizações ou exclusões.
- *Web Coverage Service (WCS)*: o termo *coverage* (“cobertura”) refere-se a um conjunto de dados em formato matricial usado para representar fenômenos com variações espaciais contínuas. O serviço WCS permite visualizar dados em formato matricial e consultar o valor numérico associado a cada pixel. Diferentemente do WFS, que manipula fenômenos geográficos discretos (vetoriais), o WCS manipula representações de fenômenos espaciais que relacionam um domínio espaço-temporal com um conjunto de propriedades. As operações suportadas por esse serviço são:
 - a) *GetCapabilities*: retorna as interfaces suportadas pelo servidor e as camadas de informação disponíveis para requisições.
 - b) *DescribeCoverage*: detalha a camada de dados matriciais requerida, apresentando informações como a localização no tempo e no espaço, os maiores e menores valores nos eixos x e y, as projeções suportadas, os formatos de saída suportados.
 - c) *GetCoverage*: retorna a camada de dados matriciais requerida, no formato especificado. A consulta pode ser feita apenas pelo retângulo envolvente ou pelo intervalo de tempo.

Tecnologias – perspectiva open source

A Figura 3 apresenta os principais componentes de uma IDE sob a perspectiva de componentes de software pela visão dos autores. Esses componentes podem ser agrupados em quatro grandes grupos: “dados” (responsável pelo gerenciamento dos dados geográficos), “servidores” (responsável pela disponibilização de serviços para publicação dos dados), “interfaces” (responsável pelos mecanismos de acesso aos dados) e “aplicações” (softwares para a interação com o usuário final). As soluções de software disponíveis atualmente possibilitam a concepção e a construção de uma infraestrutura de dados espaciais totalmente

baseada em softwares livres, e os próximos parágrafos apresentam as soluções mais difundidas em cada um desses níveis.

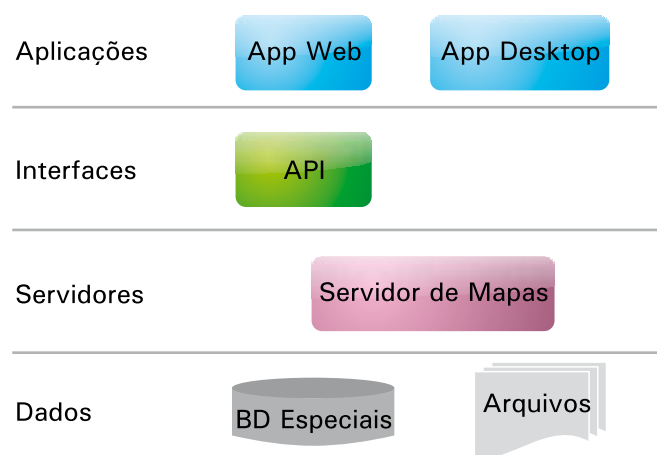


Figura 3. Componentes de uma IDE.

Fonte: elaborado pelos autores.

A camada de “dados” é a base de toda a estrutura. Um dado geográfico pode ser representado em dois tipos de estrutura de dados: matricial e vetorial. Dados matriciais (*raster*) utilizam um conjunto de pixels para representar entidades espaciais sob uma forma numérica, sendo comumente representados em arquivos nos formatos .geotiff, .jpg ou .bmp, entre outros. Os dados vetoriais identificam dados singularmente como pontos, linhas ou polígonos e podem ser representados em arquivos nos formatos *shape file* (SHP) ou *keyhole markup language* (KML), por exemplo.

Há ainda a possibilidade de armazenamento de ambos os tipos de dados em sistemas gerenciadores de banco de dados (SGBDs) espaciais, que são capazes de prover vários mecanismos para preservar a integridade dos dados e o acesso por múltiplos usuários. O SGBD *open source* mais utilizado atualmente é o PostgreSQL¹². O PostGIS¹³ foi desenvolvido pela empresa canadense *Refractions Researches* como um projeto para adicionar suporte a objetos geográficos ao PostgreSQL. Trata-se de um módulo que adiciona a capacidade de armazenamento e recuperação de entidades geográficas segundo as especificações da OGC. Além do armazenamento, esse módulo

também implementa diversas funcionalidades para análises espaciais e topológicas que estendem o próprio SQL do PostgreSQL. O código-fonte disponível para *download* está voltado para compilação em plataforma Linux/Unix, podendo ser compilado, com algumas adaptações, em qualquer plataforma. O desenvolvimento do PostGIS é contínuo, e periodicamente são acrescentadas melhorias de funcionalidades e interfaces. A Figura 4 ilustra o software pgAdmin¹⁴, que provê uma interface gráfica para a administração e a manipulação de bancos de dados PostgreSQL/PostGIS.

As camadas superiores, “servidores” e “interfaces”, são responsáveis por prover acesso aos dados para as aplicações. Essa camada viabiliza dois principais meios de acesso: direto aos dados (via APIs) e via protocolos OGC (WMS, WFS, WCS) utilizando-se servidores de mapas. A API de acesso mais difundida atualmente é a GeoTools¹⁵, que é uma biblioteca que permite o desenvolvimento de soluções de acordo com os padrões atuais OGC. É escrita na linguagem de programação Java e desenvolvida e mantida por uma comunidade de usuários. Sua arquitetura e concepção modular a tornam facilmente extensível para a adição de novas funcionalidades e favorecem sua utilização como base para o desenvolvimento de outras aplicações.

Quanto a servidores de mapas, o GeoServer¹⁶ é o mais utilizado atualmente e foi definido como servidor de mapas oficial da INDE. Trata-se de um projeto voltado para a comunidade desenvolvido, testado e apoiado por grupos diversificados e organizações de todo o mundo. O seu foco é facilitar o uso e suporte para os padrões abertos especificados pela OGC a fim de permitir o compartilhamento de informações geoespaciais de uma maneira interoperável. O GeoServer é capaz de integrar diversos repositórios de dados geográficos com alto desempenho e possui uma ferramenta de administração via *Web* que facilita a sua utilização e configuração, mostrada na Figura 5. Sua implementação é baseada em *servlets* Java (JEE), podendo ser hospedada em qualquer *servlet container*, e utiliza a API GeoTools. Entre suas funcionalidades mais sofisticadas podemos citar a

¹² Disponível em: <<http://www.postgresql.org/>> .

¹³ Disponível em: <<http://postgis.refractions.net/>> .

¹⁴ Disponível em: <<http://www.pgadmin.org/>> .

¹⁵ Disponível em: <<http://www.geotools.org/>> .

¹⁶ Disponível em: <<http://geoserver.org>> .

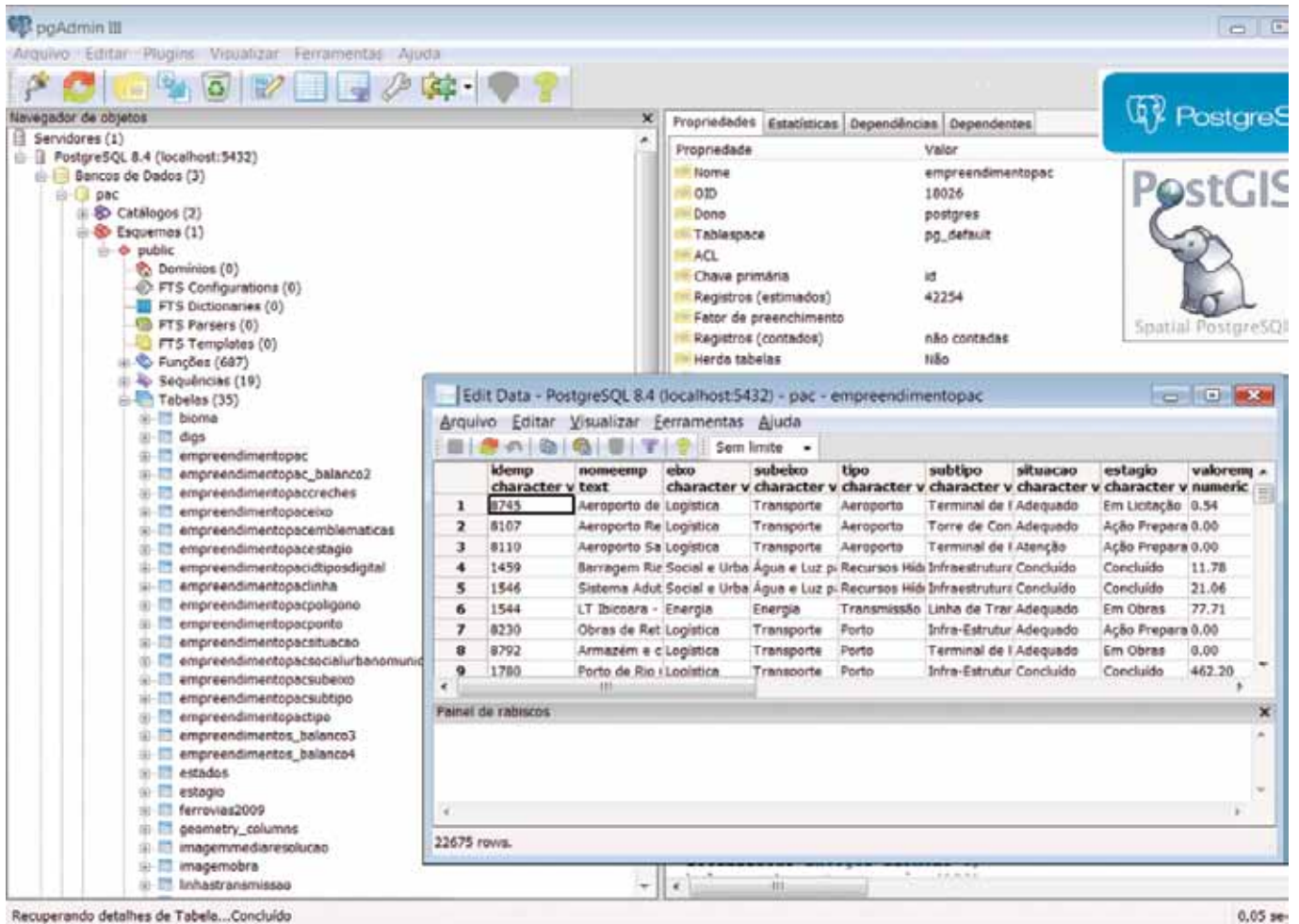


Figura 4. Interface de administração e manipulação PostgreSQL/PostGIS.
Fonte: elaborado a partir do software pgAdmin (2012).

capacidade de tratamento de projeção em tempo real, que permite unificar, de forma transparente, repositórios de dados cartográficos de diversas fontes (diferentes projeções, *datum*, fuso, dentre outros).

Na camada de “aplicações”, temos duas principais categorias a serem consideradas: aplicações de interface *desktop* e de interface *Web*. Atualmente, existem várias opções de aplicações *desktop* para acesso e manipulação de dados geográficos bem difundidas nos meios profissional e acadêmico. Algumas dessas são apresentadas a seguir.

- Spring (CÂMARA et al., 1996b): desenvolvido pela Divisão de Processamento de Imagens do Instituto Nacional de Pesquisas Espaciais (DPI/INPE) e distribuído gratuitamente para download pelo site da instituição. Seu objetivo é fornecer um ambiente unificado de geoprocessamento

e sensoriamento remoto para aplicações urbanas e ambientais e ser um mecanismo de difusão do conhecimento desenvolvido pelo INPE e seus parceiros, sob forma de novos algoritmos e metodologias. Possui funções de processamento de imagens, análise espacial, modelagem numérica de terreno e consulta a bancos de dados espaciais.

- Quantum GIS (QGIS) (HUGENTOBLE, 2008): projetado e mantido com o apoio da *Open Source Geospatial Foundation* (OSGeo), o desenvolvimento e a manutenção desse sistema são realizados por equipes de desenvolvedores espalhados pelo mundo, cada qual com um conjunto específico de responsabilidades. O QGIS é multiplataforma e permite acesso a vários tipos de dados vetoriais através da biblioteca OGR e vários formatos matriciais. Possui uma grande

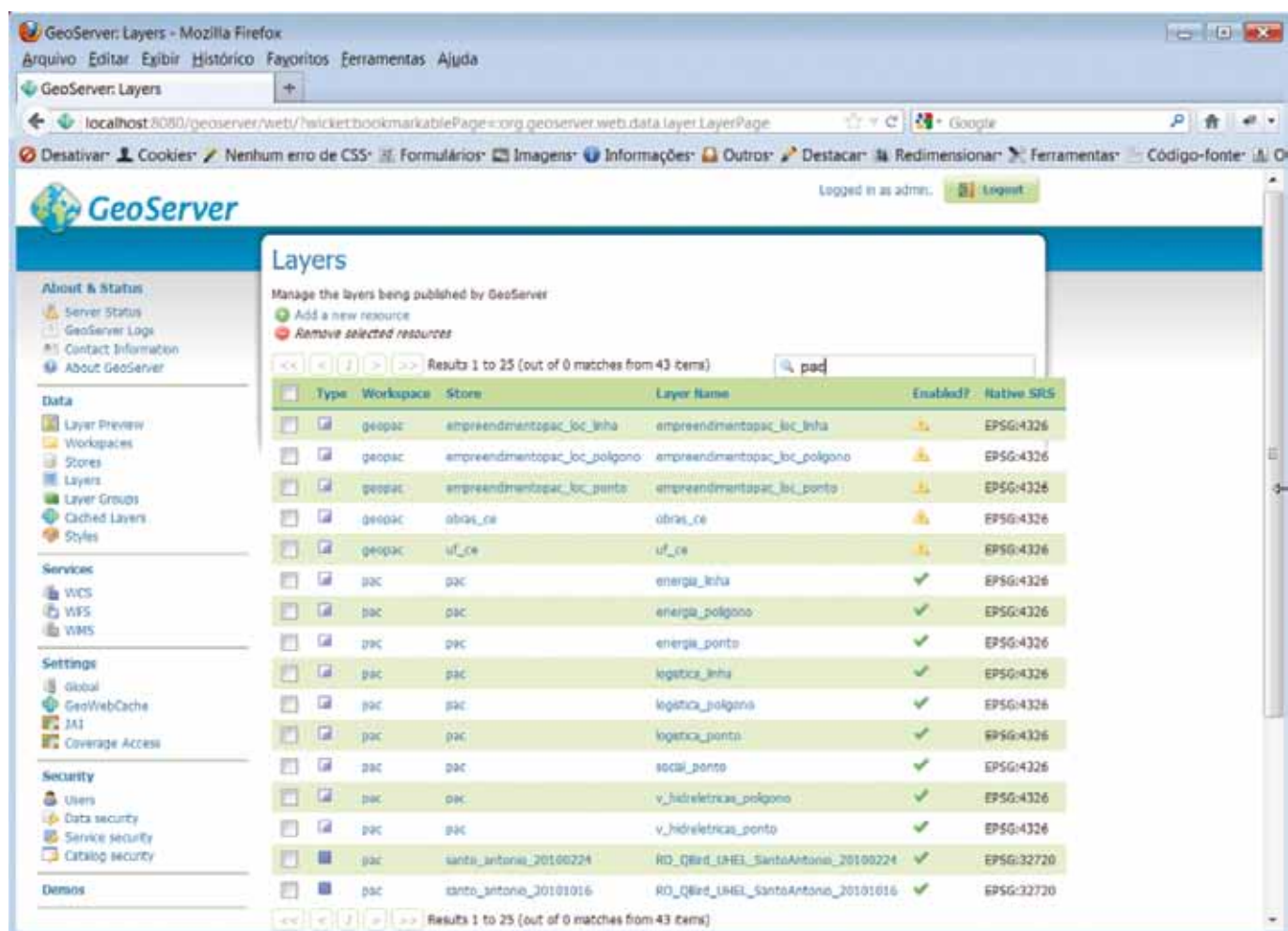


Figura 5. Interface de administração do GeoServer.
Fonte: elaborado a partir do software GeoServer (2012).

integração com o Python, uma linguagem para automatizar funções. A interface dessa aplicação é mostrada na Figura 6.

- gvSIG¹⁷: desenvolvido pela *Conselleria d'Infraestructures i Transports* (CIT) da Comunidade de Valência (Espanha) com o apoio da União Europeia, dispõe de funções para aquisição, armazenamento, gerenciamento, manipulação, processamento, exibição e publicação de dados e informações geográficas. O gvSIG possui uma arquitetura modular, permitindo a incorporação de novas funcionalidades desenvolvidas na linguagem Java. Implementa os padrões do OGC.

Para a implementação de interfaces *Web*, a principal referência é o OpenLayers¹⁸. Trata-se de uma biblioteca desenvolvida em JavaScript que permite construir páginas *Web* contendo informação geoespacial dinâmica e independente de servidor. O OpenLayers implementa métodos-padrão de acesso a dados geográficos, como os protocolos WMS e WFS, e é capaz de exibir mapas carregados a partir de qualquer fonte de dados. OpenLayers é desenvolvido e suportado por várias organizações e está constantemente à procura de patrocinadores e colaboradores para apoiar e fortalecer a comunidade. A Figura 7 mostra uma aplicação *Web* que integra bibliotecas de interface JavaScript com o OpenLayers, que acessa dados armazenados em PostgreSQL/PostGIS e publicados via GeoServer. Essa aplicação foi desenvolvida pelos autores para o atendimento de demandas de parceiros.

¹⁷ Disponível em: <<http://www.gvsig.org/Web/>>.

¹⁸ Disponível em: <<http://openlayers.org/>>.

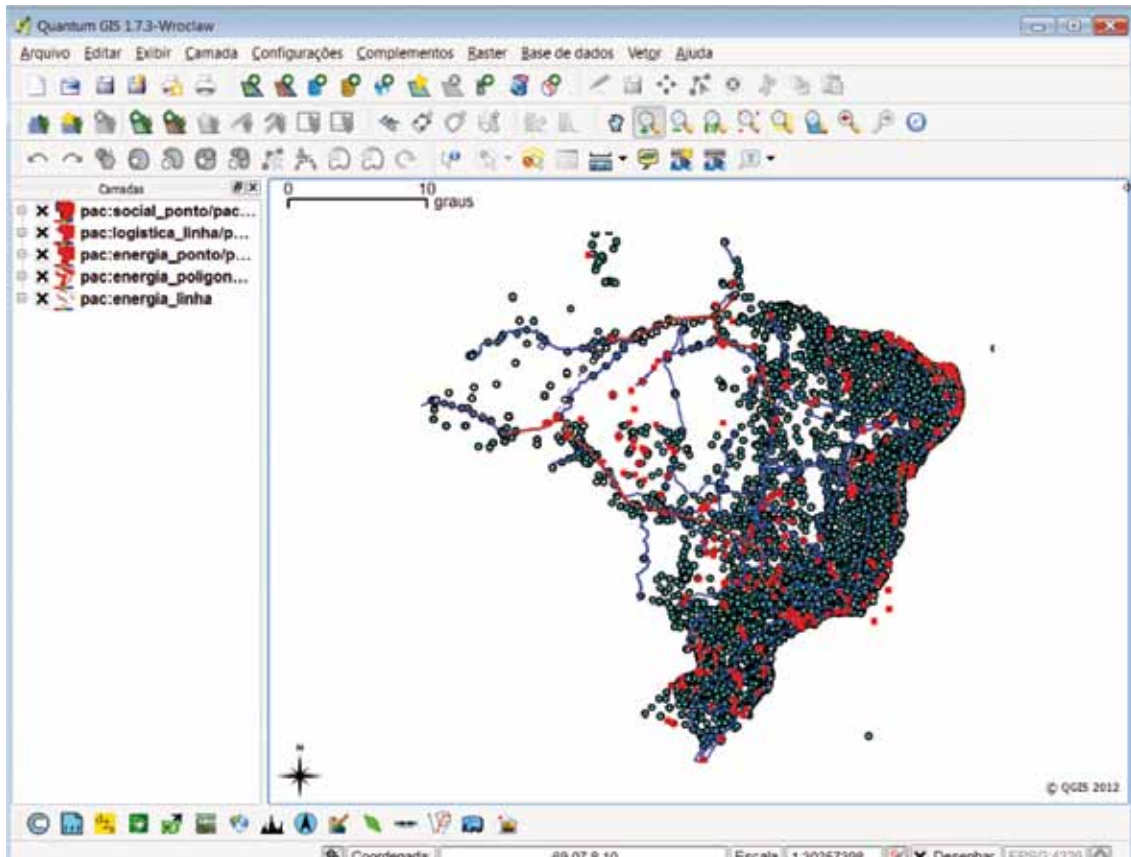


Figura 6. Interface do software *Quantum GIS*.
 Fonte: elaborado a partir do software *Quantum GIS* (2012).

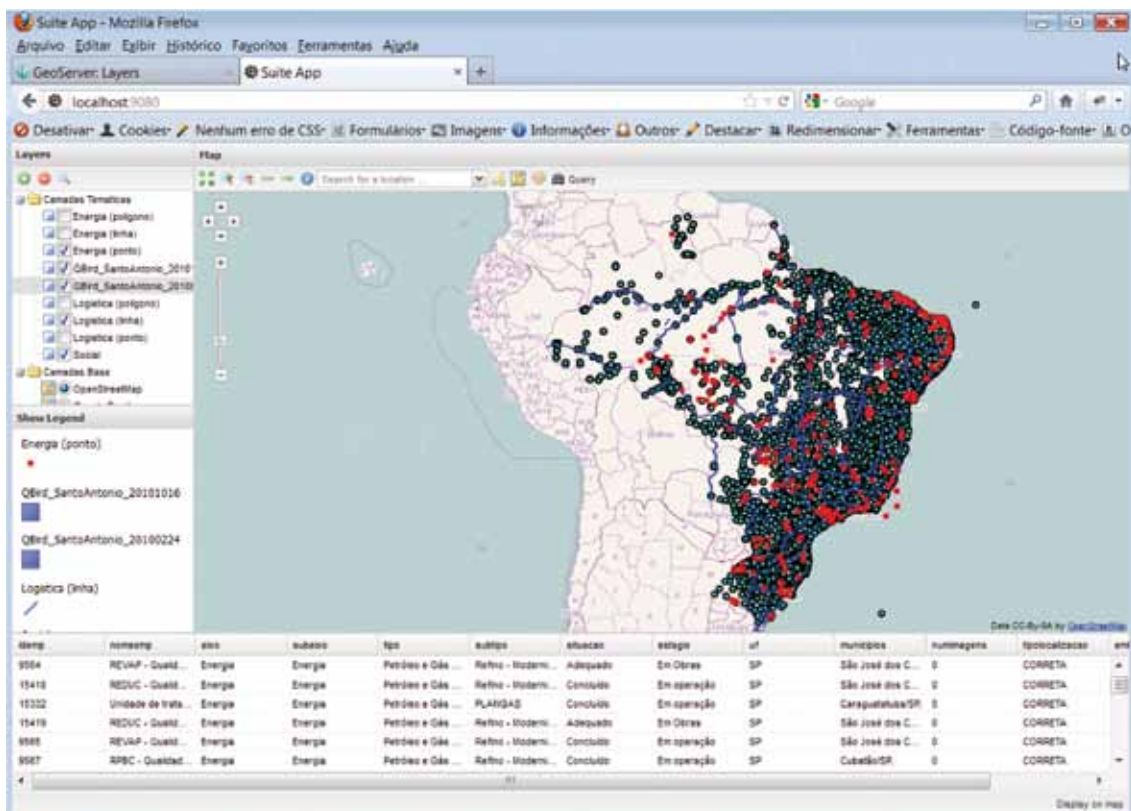


Figura 7. Aplicação *Web* em desenvolvimento pelos autores¹⁹.

¹⁹ Aplicação *Web* em desenvolvimento que integra bibliotecas de interface JavaScript e OpenLayers para acessar dados publicados pelo GeoServer via protocolos OGC (WMS,WFS).

Considerações finais

Os padrões e especificações que embasam a INDE, instituída em 2008, estão em processo de formalização e amadurecimento ao longo da última década. Grande parte deles já atingiu, atualmente, um nível de maturidade suficiente para tornar viável sua utilização em aplicações reais, com requisitos de disponibilidade e desempenho. A adoção desses padrões em iniciativas governamentais demonstra uma clara aproximação e um alinhamento de interesses entre o governo e consórcios e instituições acadêmicas, o que pode viabilizar pesquisas e resultados ainda mais expressivos.

Um tópico ainda em fase de amadurecimento é a especificação dos padrões semânticos (inclusive para dados geográficos), que atualmente ainda são alvo de pesquisa. A associação de semântica representa um dos três principais desafios a serem superados pelas novas aplicações de dados geográficos (CÂMARA et al., 2009). Há várias linhas de pesquisa para a especificação de linguagens para a representação de semântica, porém ainda é preciso avançar muito nos processos de elaboração, manipulação e integração de vocabulários e, principalmente, de associação de vocabulários a dados.

O interesse em manter a compatibilidade com os padrões de mercado também é fundamental, não só para manter o projeto tecnologicamente ativo, mas também para atrair a colaboração de profissionais e desenvolvedores da comunidade. A continuidade, atualização e sobrevivência de softwares livres, por exemplo, estão diretamente relacionadas aos incentivos e interesses da comunidade de desenvolvimento, que não é subsidiada por um grande fabricante. Manter a motivação dessas equipes é crucial para evitar a descontinuidade de um produto de software.

As soluções dos softwares livres atualmente disponíveis – SGBDs, APIs, bibliotecas e sistemas – comprovam a maturidade desses padrões e especificações e o comprometimento da comunidade em viabilizar a ampla utilização dessas tecnologias na prática. A disponibilização e o fortalecimento de uma IDE nacional, baseada em tecnologias *open source*, tornará disponível publicamente o conhecimento sobre distribuição

espacial dos principais fenômenos de interesse para o Estado e a população. Com isso, o espaço geográfico passará a ser a base referencial e a unidade integradora de informações, viabilizando a gestão territorial como instrumento de planejamento, implantação e acompanhamento das políticas públicas e de setores privados (SPADOTTO, 2012). O uso de geotecnologias, convertidas em ferramentas de sistematização de informações e conhecimento, aumenta consideravelmente a eficiência e a assertividade da gestão territorial.

A adoção e implantação da INDE nos órgãos públicos federais estão em andamento. Um exemplo disso é o grande volume de dados atualmente hospedados pelo portal de visualização da INDE²⁰, que incorpora dados de grandes instituições como o Instituto Brasileiro de Geografia e Estatística (IBGE), o Ministério do Planejamento e o Ministério de Desenvolvimento Social. Na Embrapa, esse processo tem se fortalecido com a execução de projetos como o “Modelo de Gestão da Informação Geoespacial da Embrapa”²¹, que visa fortalecer a gestão da informação geoespacial da Empresa por meio da implantação de uma estrutura de armazenamento e gerenciamento em um repositório de dados espaciais em consonância com as diretrizes da INDE. Esse projeto, iniciado em setembro de 2012, envolve 8 Unidades da Embrapa e prevê, além da catalogação de dados geoespaciais de projetos estratégicos da Embrapa, ações de capacitação de multiplicadores para a implantação e catalogação de dados e metadados geoespaciais.

Referências

ALONSO, G.; CASATI, F.; KUNO, H.; MACHIRAJU, V. **Web services: concepts, architectures and applications**. Berlin, DE: Springer, 2004. 354 p.

BOX, D.; EHNEBUSKE, D.; KAKIVAYA, G.; LAYMAN, A.; MENDELSON, N.; NIELSEN, H. F.; THATTE, S.; WINER, D. **Simple Object Access Protocol (SOAP) 1.1**. 2000. Disponível em: <<http://www.w3.org/TR/2000/NOTE-SOAP-20000508>>. Acesso em: 15 out. 2012.

BRASIL. Decreto-lei nº 6.666, de 27 de novembro de 2008. Institui, no âmbito do Poder Executivo Federal, a Infraestrutura Nacional de Dados Espaciais - INDE. **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo. Brasília, DF, 28 de nov. 2008. Seção 1, p. 57.

²⁰ Disponível em: <<http://www.visualizador.inde.gov.br/>>.

²¹ Macroprojeto SEG 05.11.11.009.00.00

- BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Padrão de metadados do governo eletrônico**. 2010a. Disponível em: <<http://www.governoeletronico.gov.br>>. Acesso em: 15 out. 2012.
- BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Padrões de interoperabilidade de governo eletrônico**. 2010b. Disponível em: <<http://www.eping.e.gov.br>>. Acesso em: 15 out. 2012.
- BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Plano de ação para implantação da INDE**. 2010c. Disponível em: <<http://www.concar.gov.br/arquivo/PlanoDeAcaoINDE.pdf>>. Acesso em: 15 out. 2012.
- BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Vocabulário controlado do governo eletrônico**. 2011a. Disponível em: <<http://vocab.e.gov.br/2011/03/vcge>>. Acesso em: 15 out. 2012.
- BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Vocabulário controlado do governo eletrônico**. 2011b. Disponível em: <<http://vocab.e.gov.br/2011/03/vcge#esquema>>. Acesso em: 15 out. 2012.
- BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Vocabulários e ontologias do governo eletrônico**. 2011c. Disponível em: <<http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padros-de-interoperabilidade/e-VoG>>. Acesso em: 15 out. 2012.
- BRAY, T.; PAOLI, J.; SPERBERG-MCQUEEN, C. M. **Extensible markup language (xml) 1.0**. 1998. Disponível em: <<http://www.w3.org/TR/1998/REC-xml-19980210>>. Acesso em: 15 out. 2012.
- CÂMARA, G.; CASANOVA, M. A.; HEMERLY, A. S.; MAGALHÃES, G. C.; MEDEIROS, C. M. B. **Anatomia de sistemas de informações geográficas**. Campinas: Unicamp, 1996a. 197 p.
- CÂMARA, G.; SOUZA, R.; FREITAS, U.; GARRIDO, J. SPRING: integrating remote sensing and GIS with object-oriented data modelling. **Computers and graphics**, New York, v. 15, n. 6, p. 13-22, 1996b.
- CÂMARA, G.; VINHAS, L.; DAVIS, C.; FONSECA, F.; CARNEIRO, T. G. S. Geographical information engineering in the 21st century. In: RESEARCH trends in geographic information science. Berlin, DE: Springer, 2009. p. 203-218.
- CASTRO, F. Falta de uma infraestrutura de dados espaciais limita pesquisa oceanográfica no Brasil. **Revista Fapesp**, São Paulo, 2012. Disponível em: <<http://agencia.fapesp.br/15472>>. Acesso em: 15 out. 2012.
- CHRISTENSEN, E.; CURBERA, F.; MEREDITH, G.; WEERAWARANA, S. **Web services description language (WSDL) 1.1**. 2001. Disponível em: <<http://www.w3c.org/TR/wsdl>>. Acesso em: 15 out. 2012.
- GEOSERVER. Disponível em: <<http://geoserver.org>>. Acesso em: 15 out. 2012.
- HOOROCKS, I. Ontologies and the semantic Web. **Communications of the ACM**, New York, v. 51, n. 12, p. 58-67, 2008.
- HUNGENTOBLE, M. Quantum GIS. In: SHEKHAR, S.; XIONG, H. (Ed.). **Encyclopedia of GIS**. Berlin, DE: Springer, 2008. p. 935-939.
- MICHAELIS, C. D.; AMES, D. P. Web feature service (WFS) and Web map service (WMS). In: SHEKHAR, S.; XIONG, H. (Ed.). **Encyclopedia of GIS**. Berlin, DE: Springer, 2008. p. 1259-1261.
- PAPAZOGLOU, M. P. Service-oriented computing: concepts, characteristics and directions. In: PROCEEDINGS OF THE FOURTH INTERNATIONAL CONFERENCE ON WEB INFORMATION SYSTEMS ENGINEERING, 2003, Washington, DC. **Anais...** Washington, DC: IEEE CS, 2003.
- PGADMIN: postgresSQL tools. Version 1.16.1. Disponível em: <<http://www.pgadmin.org/>>. Acesso em: 15 out. 2012.
- QUANTUM GIS. Version 1.8.0. Disponível em: <<http://www.qgis.org/>>. Acesso em: 15 out. 2012.
- SANTOS, R. F.; LU, C. T. Geography markup language (GML). In: SHEKHAR, S.; XIONG, H. (Ed.). **Encyclopedia of GIS**. Berlin, DE: Springer, 2008. p. 364-368.
- SPADOTTO, C. A. Ocupação agrícola: a importância da gestão territorial. **Agroanalysis**, São Paulo, v. 32, n. 9, p. 31, 2012.
- WARNEST, M. A collaboration model for national spatial data infrastructure in federated countries. 2005. 279 f. Thesis (PhD of philosophy)–The University of Melbourne, Department of Geomatics, Melbourne.
- WEIBEL, S.; KOCH, T. The dublin core metadata initiative: mission, current activities, and future directions. **D-Lib Magazine**, Reston, v. 6, n. 12. Disponível em: <<http://www.dlib.org/back2000.html>>. Acesso em: 15 out. 2012.

Comunicado Técnico, 1



Ministério da
Agricultura, Pecuária
e Abastecimento



Exemplares desta edição podem ser adquiridos na:
Embrapa Gestão Ambiental
Av. Soldado Passarinho, 303, Fazenda Chapadão.
CEP 13070-115, Capinas, SP
Fone: (19) 3211-6200
sac.sgte@embrapa.br
www.scgte.embrapa.br

1ª edição
On-line (2012)

Comitê de publicações

Presidente: *Mirian Therezinha Souza da Eira*
Secretário-Executivo: *Rosângela Galon Arruda*
Membros: *Alba Chiesse da Silva, Helena Sicoli, Ivan Sérgio Freire de Sousa, Eliane Gonçalves Gomes Assunta, Rosana Hoffman Câmara, Chang das Estrelas Wilches, Marita Féres Cardilo, Otávio Valentim Balsadi, Jeane de Oliveira Dantas*

Expediente

Supervisão editorial: *Erika do Carmo Lima Ferreira*
Revisão de texto: *Letícia Ludwig Loder*
Editoração eletrônica: *Carlos Eduardo Felice Barbeiro*