



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Uma Arquitetura de Integração de Dados Espaciais:
um Estudo dos Dados de Solos e Folhas dos Biomas
Brasileiros**

Harley Vera Olivera

Documento apresentado como requisito parcial
para a conclusão do Mestrado em Informática

Orientadora
Prof.^a Dr.^a Maristela Terto de Holanda

Brasília
2013

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Mestrado em Informática

Coordenador: Prof. Dr. Ricardo Pezzuol Jacobi

Banca examinadora composta por:

Prof.^a Dr.^a Maristela Terto de Holanda (Orientadora) — CIC/UnB
Prof.^a Dr.^a Maria Emília Machado Telles Walter — CIC/UnB
Prof. Dr. Henrique Llacer Roig — IG/UnB

CIP — Catalogação Internacional na Publicação

Olivera, Harley Vera.

Uma Arquitetura de Integração de Dados Espaciais: um Estudo dos
Dados de Solos e Folhas dos Biomas Brasileiros / Harley Vera Olivera.
Brasília : UnB, 2013.

68 p. : il. ; 29,5 cm.

Dissertação (Mestrado) — Universidade de Brasília, Brasília, 2013.

1. Integração de dados, 2. Banco de Dados Geográficos, 3. Web
Service, 4. Modelagem de Dados geoespaciais, 5. Dados Geoespaciais,
6. Arquitetura Orientada a Serviços, 7. sistemas de Informação
Geográficas.

CDU 004

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil

Dedicatória

Dedico este trabalho aos meus pais por ser a base fundamental em tudo que eu sou, em toda a minha educação, tanto acadêmica, como da vida, pelos seus conselhos, seus valores, pelo seu apoio incondicional perfeitamente mantido através do tempo, mas, acima de tudo, pelo seu amor.

A meus irmãos e família, que com muito carinho e apoio, não mediram esforços e sacrifícios para que eu chegasse até esta etapa de minha vida.

A minha orientadora Prof.^a Dr.^a Maristela Terto de Holanda exemplo de disciplina e organização, pela enorme paciência que teve comigo, pelas muitas vezes que deixou de ser a orientadora para virar a mãe que cuida do filho, pelo seu apoio e motivação para a culminação do mestrado.

Agradecimentos

Quero Agradecer a todas as pessoas que me apoiaram desde o primeiro momento que tomei a decisão de começar o mestrado, quando cheguei nesse novo país, e no decorrer dos estudos até a culminação do mestrado. Sem o apoio deles nada do que consegui ate hoje houvesse sido possível.

A decisão de fazer estudos fora do país de origem não é simples, por isso quero agradecer a minha família pelo seu apoio incondicional, aos meus pais e irmãos, que sacrificaram a presença do seu filho e irmão somente por ter a satisfação de ver-me superado numa nova etapa da minha vida.

A todos os professores do departamento de Computação, que foram tão importantes na minha vida acadêmica e no desenvolvimento deste trabalho, de maneira especial a minha orientadora Prof. Maristela Terto de Holanda, pelas aulas, pelas sugestões, pelos conselhos e dicas de pesquisa, pelo material emprestado, pela paciência que teve comigo, pela participação e pela ajuda incondicional, quem com seus conhecimentos e experiência soube me encaminhar no mestrado. Torço pra que esta parceria continue por muitos anos. Ao professor Henrique Llacer Roig e a professora Maria Emília Machado Telles Walter que são parte da minha banca, pela sua presença, suas sugestões e contribuições para com meu trabalho.

Agradeço também aos meus colegas do mestrado com quem compartilhei bons momentos durante meus estudos, de maneira especial a todos que me ajudaram nos momentos de duvidas e dificuldades, muito obrigado Ruben Cruz, Henrique Freitas, Zheng Jianya, Yang Liu, Wosley Arruda, Liliane Nascimento. Agradeço também ao Eder Santana e Diego Villacorta pela imprescindível ajuda no trabalho. Aos doutores do Departamento de Ecologia da UnB Fred Takahashi e Alexandre de Siqueira pela ajuda desinteressada em cada momento.

Aos meus amigos peruanos, com quem foi mais fácil matar as saudades de nossas famílias e nosso país, pela companhia nos bons e maus momentos, em especial aos meus amigos Helard Becerra, Ruben Cruz, Maria Laura Chavez, Jose Luis Soncco e Juan Cruz.

Resumo

Um sistema de informação geográfico (SIG) pode ser desenvolvido utilizando bancos de dados geográficos distribuídos, mas novos desafios para a gestão de dados ambientais requerem interoperabilidade, integração e compartilhamento de dados. O presente trabalho propõe modelos de dados para Solos e Folhas do Bioma Cerrado Brasileiro e uma arquitetura para integração de dados e seus respectivos serviços web. Esses serviços são importantes para integração e compartilhamento dos dados geográficos. O trabalho desenvolvido foi baseado nas recomendações da Infraestrutura Nacional de Dados Espaciais (INDE), *web services* seguindo os padrões da *Open Geospatial Consortium* (OGC) e a utilização de software livre como PostgreSQL, GeoServer, OpenScales e XGIS no desenvolvimento do sistema de informação geográfica.

O SIG permite visualizar dados geográficos relativos a solos e vegetação presentes no território brasileiro. Isto se faz necessário, para facilitar a utilização pelos pesquisadores do Departamento de Ecologia da UnB dos seus dados obtidos de dissertações, teses, artigos, etc. que são armazenados em planilhas Excel. Atualmente, a análise desses dados é complexa devido à dificuldade de manuseá-lo através das planilhas. Para isso, foi realizado um estudo e análise das planilhas para obter os modelos de dados relacionais para solos e vegetação. A partir dessa modelagem, com ajuda de uma ferramenta de extração de dados Geokettle foi possível a implementação do banco de dados previamente implementado no PostgreSQL com sua extensão PostGIS.

Com os bancos de dados prontos, foram disponibilizados mapas de pontos de solos e vegetação mediante uma serviços web para dados geográficos, a ferramenta utilizada foi o Geoserver. Finalmente foi implementado o SIG para visualização e análise desses dados, utilizando tecnologias como o *framework* de mapas OpenScales e o *framework* XGIS para geração da interface gráfica.

Palavras-chave: Integração de dados, Banco de Dados Geográficos, Web Service, Modelagem de Dados geoespaciais, Dados Geoespaciais, Arquitetura Orientada a Serviços, sistemas de Informação Geográficas.

Abstract

A geographic information system (GIS) can be developed using distributed geographic databases, but new challenges for the management of environmental data require interoperability, integration and data sharing. This work proposes data models for soils and leaves of the Brazilian Cerrado, an architecture for data integration and their web services. These services are important for integration and sharing of spatial data. The work was based on the "Comissão da Infraestrutura Nacional de Dados Espaciais" (CINDE) recommendations and follows the web services standards of the Open Geospatial Consortium (OGC). The use of open source software was taken into account for the development of this geographic information system.

The GIS allows to visualize geographic data related to soils and vegetation present in the Brazilian territory. This is necessary to facilitate the use by researchers in the Department of Ecology at UnB for their data obtained from dissertations, theses, articles, etc. that are stored in Excel spreadsheets. Currently, the analysis of these data is complicated due to the difficulty of handling it through spreadsheets. For this, a study and analysis of spreadsheets was made to get relational data models for soils and vegetation. From this modeling, with the help of a tool for extracting data named Geokettle was possible to implement the database previously implemented in PostgreSQL with its extension PostGIS.

With databases ready data were available by maps of points of soils and vegetation through a web services for spatial data, the tool used was the Geoserver. Finally we implemented the GIS for visualization and data analysis, using technologies such as framework of maps OpenScale and XGIS for generating graphical interface.

Keywords: Data integration, Geospatial database, Web Service, Modeling Geospatial Data, Service Oriented Architecture, Geographical Information System.

Sumário

1	Introdução	1
1.1	Contextualização	1
1.2	Objetivos	2
1.2.1	Objetivo Geral	2
1.2.2	Objetivos Específicos	2
1.3	Estrutura do Trabalho	2
2	Sistemas de Informação Geográfica	3
2.1	Paradigma dos Quatro Universos	3
2.1.1	O Universo Ontológico	3
2.1.2	O Universo Formal	4
2.1.3	O Universo Estrutural	7
2.1.4	O Universo de Implementação	8
2.2	Dados Geográficos	8
2.3	Banco de Dados Geográficos	9
2.4	Modelagens Conceituais de Dados Geográficos	9
2.4.1	Técnica de Modelagem de Objetos para Aplicações Geográficas	10
2.4.2	Infraestrutura Nacional de Dados Espaciais - INDE	13
2.4.3	Documento das Especificações Técnicas para a Estruturação de Dados Geoespaciais e Vetoriais	16
2.4.4	<i>Open Geospatial Consortium</i> (OGC)	17
2.5	Arquiteturas Orientadas a Serviços	18
2.6	Sistema de Informação Geográfico para Web	23
2.6.1	Arquitetura de um SIG	23
2.6.2	Necessidades de Aplicações de Geoprocessamento	24
3	Modelos de dados e <i>Web Services</i>	26
3.1	Ferramenta de Modelagem de Banco de Dados	26
3.2	Sistema Gerenciador de Banco de Dados	27
3.3	Extensões Espaciais	27
3.4	Servidor de Serviços Geográficos	27
3.5	Quantum GIS	28
3.6	GeoKettle	29
3.7	OpenScales	30
3.8	XGIS	30
3.9	Arquitetura Proposta para Integração de Dados Espaciais	31

4	Resultados	33
4.1	Modelo de Banco de Dados Geográfico	33
4.1.1	Temas Relevantes Abordados no Processo de Modelagem	33
4.1.2	Amostras de Vegetação e Solos	33
4.1.3	Modelos e Especificações para Modelagem de Banco de Dados Geográficos	36
4.2	Relacionamentos dos Modelos de Dados	38
4.3	Construção dos Bancos de Dados	45
4.3.1	Criação de Modelo Lógico	45
4.3.2	Criação de Modelo Físico	48
4.3.3	Criação do <i>Script</i> de Construção do Banco de Dados	48
4.4	Carga de Dados nos Bancos de Dados Geográficos	48
4.5	Modelo de Dados Final	50
4.6	Demonstração do Sistema de Informação Geográfica	50
5	Conclusões	54
5.1	Contribuições	55
5.2	Trabalhos Futuros	55
	Referências	56

Lista de Figuras

2.1	Paradigma dos Quatro Universos [1].	3
2.2	Dualidade entre Espaço Absoluto e Espaço Relativo [1]	6
2.3	Exemplo de Geo-campo e Conjunto de Geo-objetos [1]	7
2.4	Representação das estruturas vetoriais [1]	7
2.5	Formas Derivados dos Dados Espaciais a partir dos Básicos [1].	8
2.6	Notação Gráfica para Classes no modelo OMT-G [2].	11
2.7	Geo-campos [2].	12
2.8	Geo-objetos (a) com geometria, (b) com geometria e topologia[2].	12
2.9	Relacionamentos [1].	13
2.10	Pilares de uma IDE [1].	15
2.11	Categorias de informação da CONCAR [3].	16
2.12	Arquitetura SOA-SDI de Quatro Camadas [4].	20
2.13	Arquitetura Orientada a Serviços Baseada em Três Níveis [5].	22
2.14	Arquitetura Proposta no Plano de Ação da INDE [6].	23
2.15	Arquitetura de Sistemas de Informações Geográficas [1].	24
3.1	Tela com as camadas existentes no GeoServer.	28
3.2	Visualização dos pontos de solos e vegetação no mapa dos Biomas de Brasil. Fonte propria.	29
3.3	Ferramenta de extração, carregamento e transformação de dados Geokettle [25].	30
3.4	Arquitetura Proposta para Integração de Dados Geoespaciais.	31
4.1	Dados de Localização Geográfica da Planilha Solos.	34
4.2	Detalhes de desenvolvimento da classe Localizacao geografica.	37
4.3	Relações Geográficas e não Geográficas no Modelo de dados do Solo.	38
4.4	Relações Geográficas e não Geográficas, Modelo de dados da Vegetação.	39
4.5	Fragmento Modelo de Dados vegetação.	39
4.6	Fragmento Modelo de Dados de Solos.	40
4.7	Especificações dos Atributos Entidade Temperatura anual Modelo de dados da Vegetação.	41
4.8	Integração Entidades Horizonte - Quimica amostra Modelo de Dados do Solo.	42
4.9	Entidades <i>Fosforo, Nitrogenio, Carbono Integradas a Entidade Quimica</i> Modelo de Dados Solos.	42
4.10	Integração Entidades <i>Referencias e Amostra solos</i> Modelo de Dados Solos.	43
4.11	Entidade <i>Mudancas uso terra</i> integrada a <i>Localização geográfica</i>	43

4.12	Integrações com Entidades Simples com a Entidade Central <i>Amostra vegetacao</i>	44
4.13	Integração entidade <i>Descricao Taxonomica</i> e <i>Amostra Vegetacao</i>	45
4.14	Modelo de Dados Lógico do Solo.	46
4.15	Modelo de Dados Lógico da Vegetação.	47
4.16	Processo de Extração de Dados da Entidade <i>Localizacao geografica</i> e Carregamento no Banco de Dados dos solos.	49
4.17	Tela Inicial do SIG.	51
4.18	Janela que exibe as espécies para cada ponto.	52
4.19	Funções <i>Add Layers Dynamically</i> , <i>Total Extension</i> e <i>Overview</i>	52
4.20	Função <i>Export</i> , utilizada para exportar camadas.	53
4.21	Ferramenta com funções de medições sobre o mapa.	53

Glossário

CONCAR	Comissão Nacional de Cartografia, 1
GIS	Geographical Information System, vii
GML	<i>Geography Markup Language</i> , 4
IDE	Infraestrutura de Dados Espaciais, 13
INDE	Infraestrutura Nacional de Dados Espaciais, 1
LGPL	<i>Lesser General Public License</i> , 29
OGC	<i>Open Geospatial Consortium</i> , 4
OLTP	<i>Online Transaction Processing</i> ou Processamento de Transações em Tempo Real, 29
RIA	<i>Rich Internet Application.</i> , 29
SGBD	Sistemas Gerenciadores de Bancos de Dados, 9
SIG	Sistema de Informação Geográfica, 1
SIRGAS	Sistema de Referência Geocêntrico para as Américas 2000, 28
SOA	<i>Services Oriented Architecture</i> , 1
TIN	<i>Triangulated Irregular Network</i> , 11

Capítulo 1

Introdução

1.1 Contextualização

Diariamente no mundo, informações espaciais são utilizadas. Estas informações são todas aquelas referenciadas no espaço, como, as imagens da terra obtidas por satélites, os dados cartográficos e topográficos que representam um território, as redes de ferrovias, a distribuição da população e suas variáveis socioeconômicas. Esses dados espaciais estão disponíveis em arquivos e em bancos de dados. Estas informações em geral são encontradas de forma isolada, em diferentes formatos, em sistemas que geralmente não conversam entre si e servindo apenas aos objetivos da instituição que a desenvolveu.

Com dados espaciais fisicamente distribuídos em locais diferentes surgiu o desafio da interoperabilidade entre os bancos de dados e a integração de dados. A utilização de *web services* tem sido uma das alternativas para permitir a disponibilização dos dados de um determinado banco de dados, com a finalidade de ter acesso à informação e processos remotos. Para definição de padrões a serem aplicados com *web services*, a *Open Geospatial Consortium* (OGC) especifica um conjunto de formatos que devem ser utilizados para permitir o acesso aos dados geográficos. A OGC é uma organização voluntária internacional com mais de 280 organizações comerciais, governamentais, não-lucrativas e instituições de pesquisa do mundo todo.

Os padrões estabelecidos pelo OGC propõem arquitetura de serviço web para dar suporte à interoperabilidade de dados em desenvolvimento de sistemas de informação geográfica (SIG). Uma arquitetura orientada a serviços (SOA) é definida como uma arquitetura para interoperabilidade de sistemas por meio de conjunto de interfaces de serviços fracamente acoplados, onde os serviços não necessitam de detalhes técnicos da plataforma dos outros serviços para a troca de informações serem realizadas [7].

No Brasil, a Comissão Nacional de Cartografia (CONCAR), órgão colegiado do Ministério do Planejamento, Orçamento e Gestão, mediante comitês especializados elaborou as especificações técnicas para estruturação de dados geoespaciais vetoriais, publicadas no documento “Plano de Ação para Implantação da INDE (Infraestrutura Nacional de Dados Espaciais)”, pelo qual toda instituição pública e privada no Brasil com intenção de desenvolver SIG deve seguir as normas e especificações do CONCAR [6]. Os modelos de dados apresentados na documentação da INDE explora um conjunto básico de dados, nos quais a parte da Ecologia, especificamente solos e folhas, não estão contemplados.

Desta forma, um dos objetivos dessa dissertação é apresentar uma proposta de modelos de dados para solos e folhas dos Biomas Brasileiros.

O estudo de caso utilizado para validação do modelo proposto foi aplicado no projeto Respostas às mudanças climáticas: uma síntese sobre a estrutura e funções dos ecossistemas de Cerrado e da Amazônia brasileira do Departamento de Ecologia da UnB, o qual tem como objetivo tratar informações a respeito da variabilidade do clima e impactos das oscilações climáticas de processos funcionais, como uso e disponibilidade hídrica, fluxos de carbono, nitrogênio e energia [8].

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo do trabalho é implementar um Sistema de Informação Geográfico baseado numa arquitetura orientada a serviços para a integração e compartilhamento dos dados de vegetação e solos dos biomas brasileiros.

1.2.2 Objetivos Específicos

Pelo caráter multidisciplinar da pesquisa, e no intuito de atingir o objetivo geral, foram definidos alguns objetivos específicos:

- Definição da proposta de integração de dados.
- Implementação dos *web services* baseados nos modelos de dados para a utilização em uma arquitetura de dados baseadas em serviço.
- Extender o documento das especificações técnicas para estruturação de dados geoespaciais digitais vetoriais para amostras de solos e vegetação dos biomas brasileiros.

1.3 Estrutura do Trabalho

Esse documento está dividido nas seções apresentadas a seguir:

- O Capítulo 2 apresenta os conceitos de Sistemas de Informação Geográficos utilizados na pesquisa.
- O Capítulo 3 propõe os modelos de dados e *web services* desenvolvidos na pesquisa.
- O capítulo 4 discute os resultados obtidos durante a pesquisa.
- Finalmente, no Capítulo 5 são apresentadas as conclusões do trabalho.

Capítulo 2

Sistemas de Informação Geográfica

Neste capítulo são apresentadas conceitos básicos, usados nesta pesquisa, e está organizado da seguinte forma. A Seção 2.1 apresenta o paradigma dos quatro universos. A Seção 2.2 apresenta fundamentos teóricos sobre dados geográficos. A Seção 2.3 apresenta conceitos fundamentais sobre banco de dados geográficos. Na Seção 2.4 é apresentado o conceito de modelagem de dados geográficos. A Seção 2.5 apresenta as arquiteturas orientadas a serviços para compartilhamento de dados geográficos. A Seção 2.6 trata sobre sistemas de informação geográficos.

2.1 Paradigma dos Quatro Universos

Geoinformação é toda informação passível de espacialização, ou seja, tem algum tipo de vínculo geográfico que permite sua localização. Este pode ser um ponto, um endereço, um território, entre outros. O problema da geoinformação é a representação computacional do espaço geográfico, Para auxiliar nesse problema foi definido o paradigma dos quatro universos em [1]. Nesse paradigma são descritos quatro passos para transformar os dados obtidos no mundo real em informações relevantes representadas computacionalmente. Cada passo é representado por um universo: o universo Ontológico, Formal, Estrutural e de Implementação, como se apresenta na Figura 2.1.

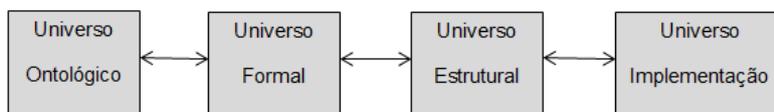


Figura 2.1: Paradigma dos Quatro Universos [1].

2.1.1 O Universo Ontológico

Ontologia é o campo da filosofia cujo objetivo é descrever os tipos e estruturas de entidades, eventos, processos e relações que existem no mundo real [9]. O interesse recente por ontologias em sistemas de informação é por causa da interoperabilidade e compartilhamento de informação de forma eficiente para um público cada vez mais interdisciplinar.

No contexto da computação, uma ontologia é um modelo de dados que representa um conjunto de conceitos dentro de um domínio e os relacionamentos entre estes [10].

Um sistema de informação é concebido como um mecanismo de comunicação entre duas partes: o produtor e o usuário. Para que o sistema funcione é preciso que os dois atores concordem com os conceitos utilizados no sistema sendo assim o sucesso dependerá da existência de uma comunidade disposta a compartilhar as definições utilizadas para construí-lo. Por exemplo, pesquisadores de amostras de solos e vegetação precisam definir conceitos relativos às amostras que sejam comuns aos dois tipos de coletas, por exemplo, o tipo de solo onde a amostra foi coletada, a caracterização do local, a temperatura anual, precipitação anual, já que elas poderiam ser interpretadas de distintas formas por cada um dos pesquisadores. Além do conceito utilizado, é preciso que o usuário compreenda como será representado esse conceito e entenda as limitações do seu uso [1]. Deste modo, para começar um projeto de desenvolvimento de um sistema de informação, é necessário primeiramente definir as entidades a serem representadas e os conceitos que se aplicam a elas, e se quisermos que esses conceitos sejam compartilhados por uma comunidade interdisciplinar, é fundamental que os conceitos utilizados sejam devidamente explicitados.

Para os dados geográficos, uma geo-ontologia tem dois tipos básicos de conceitos: conceitos que correspondem aos fenômenos físicos do mundo real e conceitos criados para representar entidades sociais e institucionais. O primeiro refere-se aos conceitos que correspondem aos fenômenos físicos do mundo real. O segundo corresponde aos conceitos criados para retratar entidades sociais e institucionais [1].

A maior parte dos sistemas de informação atualmente utilizados não possuem ontologias de aplicação associados a eles, o que reduz a capacidade de compartilhar informações [1]. Entretanto para sistemas de informação geográficas (SIG's) e dados geográficos em geral, o consórcio *Open Geospatial Consortium* (OGC) propôs o formato *Geography Markup Language*(GML) para descrições de ontologias [11].

2.1.2 O Universo Formal

O universo formal representa um componente intermediário entre o universo ontológico e as estruturas de dados e algoritmos computacionais. Como o computador trabalha com bits e estruturas matemáticas, fica difícil fazer uma interpretação dos conceitos informais da ontologia de aplicação para estruturas de dados, sem gerar algumas inconsistências. Então para fazer a interpretação sem inconsistências, o universo formal busca associar da forma mais abrangente possível, um conjunto de entidades lógicas aos diferentes conjuntos de conceitos da ontologia da aplicação. O universo formal também especifica como associar valores aos diferentes conceitos apresentados, ou seja, como medir os dados do mundo real [1].

Atributos de Dados Geográficos: Teoria da Medida

Para representar dados geográficos no computador devemos descrever a sua variação no tempo e espaço e saber o valor dela num certo momento, para isso é preciso compreender o processo de mensuração da realidade de forma consistente. O processo de medida consiste em associar números ou símbolos a diferentes ocorrências de um mesmo atributo. Por exemplo, para medição do nível de argila em solos, os pesquisadores coletam amostras em diferentes partes do solo, cada amostra de argila que foi coletada dará um valor de medida

diferente. Essas diferenças de atribuição de valor são chamadas escalas de medidas. A referência geral sobre escalas de medidas pode ser encontrada no trabalho de Stevens [12] que propõe quatro escalas de mensuração: nominal, ordinal, intervalo e razão.

Os níveis ordinal e nominal são temáticos, porque a cada medida é atribuído um número ou nome associando a observação a um tema ou classe. A escala nominal classifica objetos em classes distintas sem ordem inerente, como rótulos que podem ser quaisquer símbolos. As possíveis relações entre os valores são identidade ($a = b$) e dessemelhança ($a \neq b$). Um exemplo é a vegetação predominante, com rótulos como “campo sujo”, “campo limpo”, “pradaria”, etc. A escala ordinal introduz uma ideia de ordenação, caracterizando os objetos em classes distintas que possuem uma ordem natural sendo relevante apenas saber se um elemento é maior que o outro. Um exemplo de ordinalidade é o agrupamento em escalas do tipo 1 – ruim, 2 – bom, 3 – ótimo. A distância definida entre os elementos não é significativa.

Por outro lado, as medidas temáticas não estão associadas à magnitude do fenômeno e são utilizadas quando o estudo necessita de uma descrição mais detalhada que permita medir a magnitude do fenômeno e comparar intervalo e ordem de grandeza entre eventos, então são utilizadas as escalas de medida baseadas em números reais.

As escalas utilizadas nessa monografia foram principalmente as temáticas para a representação dos dados. A escala temática foi escolhida visto que o objetivo do trabalho foi representar dados geográficos contidos nas planilhas do projeto sobre amostras de vegetação e amostras de solos.

Espaço Absoluto e Espaço Relativo

Existe a possibilidade de representação no computador da localização dos objetos no espaço ou o posicionamento relativo entre eles como ilustrado na Figura 2.2, onde se mostra essa distinção através do mapa da cidade de São Paulo e seus distritos. Nesta figura mostramos à esquerda o mapa da cidade de São Paulo e seus distritos identificados por suas fronteiras, neste caso trata-se de uma representação no espaço absoluto. Do lado direito se mostra uma representação no espaço relativo, pois mostramos somente um grafo com as conexões dos distritos que formam uma rede (mostrando o mapa de fundo apenas como uma referência). Neste modelo a localização exata de cada distrito não é armazenada, pois a rede só captura as relações de adjacência [1].

Escolher qual tipo de representação será usado é de suma importância no projeto de um SIG. Esta escolha depende basicamente do tipo de análise a ser realizada pelo sistema. Representações em espaço absoluto são usualmente utilizadas quando as consultas espaciais a serem realizadas envolvem entidades de tipos diferentes. Por exemplo, quando se quer saber quantos rios cruzam um determinado município. Quando os procedimentos de análise envolvem apenas as relações de conectividade, como, quando se precisa saber o caminho mais curto entre dois distritos.

Como um dos objetivos do projeto é a visualização das localizações geográficas onde foram coletados os dados das amostras dos solos e vegetação, foram utilizadas ambas as representações. A representação no espaço absoluto foi utilizada para dividir o mapa mundi em países, assim como o mapa de Brasil em biomas e estados. E para representar os caminhos, as estradas, os rios que passam perto de cada ponto de amostra é necessário utilizar a representação no espaço relativo.

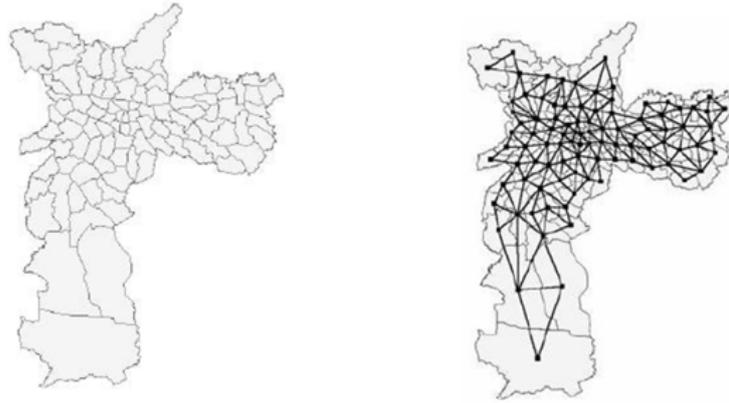


Figura 2.2: Dualidade entre Espaço Absoluto e Espaço Relativo [1]

Modelos no Espaço Absoluto: Geo-Campos e Geo-Objetos

No modelo de espaço absoluto existem dois tipos de modelos formais para entidades geográficas: o geo-campo e o geo-objeto [1] mostrado na Figura 2.3. O modelo de geo-campo enxerga o espaço geográfico como sendo contínuo, onde não se podem distinguir exatamente as partes que compõem um campo. Como exemplo, um mapa de vegetação associa a cada ponto do mapa um tipo específico de cobertura vegetal diferente ou a foto via satélite de uma parte determinada de uma cidade. Por outro lado, o modelo de geo-objetos representa o espaço geográfico como uma coleção de entidades distintas e identificáveis, onde cada entidade é definida por uma fronteira fechada. Como exemplo, temos o loteamento de um terreno [1].

Uma das características dos geo-campos é que, ao se dividir em dois, tem-se dois geo-campos com as mesmas propriedades do anterior, apesar de provavelmente ter atributos diferentes do anterior, pois cada parte do geo-campo pode ser identificada pela sua localização. Por outro lado, ao dividir um geo-objeto, esse perde as características que tinha quando era uma entidade só, por abranger entidades que podem não ter mais uma fronteira bem definida. Por exemplo, a divisão de um município em dois. Ao não ser que sejam definidas suas novas fronteiras e códigos, os dois novos geo-campos não poderão existir.

A diferença essencial entre um geo-campo e um geo-objeto está determinada pela fronteira. A fronteira de um geo-campo é uma divisão arbitrária relacionada apenas com nossa capacidade de medida. Por contraste, um geo-objeto é essencialmente definido por sua fronteira, que a separa do mundo exterior, ele não pode ser dividida e manter suas propriedades essenciais [1].

Todos os dados geográficos utilizados nesse projeto são geo-objetos exceto as amostras, tanto os coletados pelos pesquisadores do Departamento de Ecologia, como coordenadas de longitude e latitude de um ponto no planeta, como aqueles disponibilizados por órgãos públicos e utilizados no sistema implementado como suporte, como mapas do mundo, do Brasil e dos biomas.

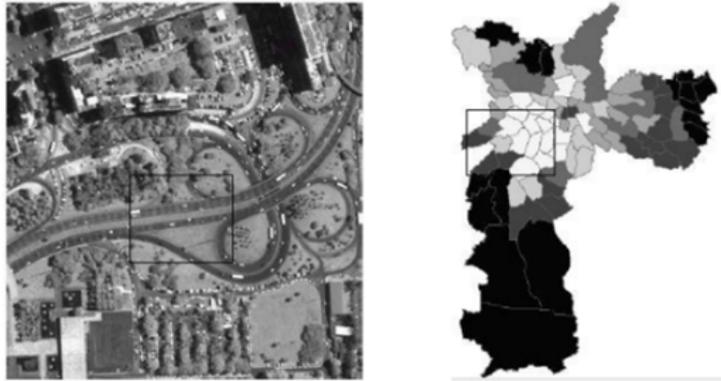


Figura 2.3: Exemplo de Geo-campo e Conjunto de Geo-objetos [1]

2.1.3 O Universo Estrutural

O universo estrutural é onde os diversos modelos formais serão representados por estruturas de dados. As estruturas de dados utilizados em banco de dados geográficos podem ser divididas em duas grandes classes: estruturas vetoriais e estruturas matriciais [1]. Nesta monografia se aprofundará em dados vetoriais notando-se que não foi utilizada nenhuma estrutura matricial para representar algum tipo de dado.

As estruturas de dados vetoriais são utilizadas para representar as coordenadas das fronteiras de dados espaciais através de três formas básicas: pontos, linhas e áreas definidas por suas coordenadas cartesianas como se mostra na Figura 2.4.

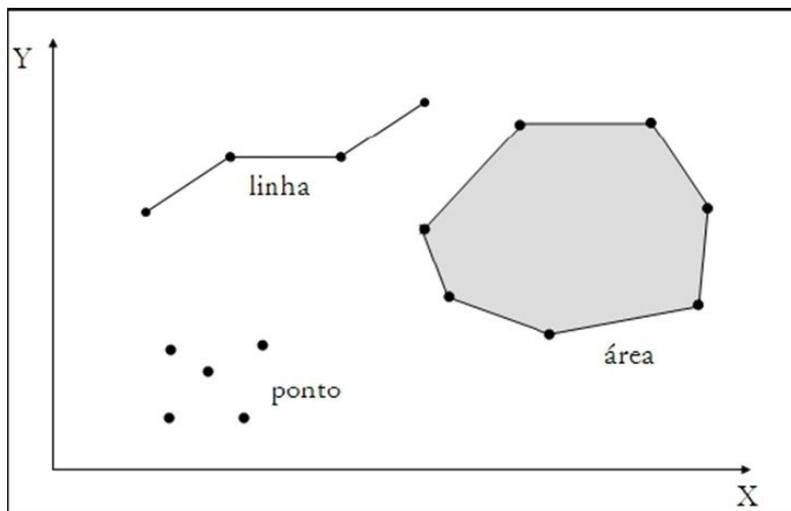


Figura 2.4: Representação das estruturas vetoriais [1]

Um ponto é um par ordenado (x, y) de coordenadas espaciais. Por exemplo, as amostras de vegetação nos biomas de Brasil [1].

Uma linha é um conjunto de pontos conectados. Como por exemplo, as estradas que conectam dois estados [1].

Um polígono é uma região de um plano que é limitada por uma ou mais linhas conectadas formando assim uma área fechada. A fronteira do polígono divide o plano em duas regiões: o interior e o exterior. Os polígonos são usados para representar unidades de dados geográficos espaciais individuais [1].

2.1.4 O Universo de Implementação

No quarto e último universo, são tomadas as decisões de programação e que podem admitir grandes variações. Estas decisões podem levar em conta a disponibilidade de algoritmos, o tipo de usuário alvo, a disponibilidade de hardware, etc.

2.2 Dados Geográficos

Os dados espaciais são definidos como aqueles que se distinguem essencialmente pelo componente espacial, que associa a cada entidade ou fenômeno uma localização na terra, traduzidas por sistemas geodésicos de referência, por exemplo, *World Geodetic system* (WGS84)[6].

As estruturas vetoriais são utilizadas para representar as coordenadas das fronteiras de cada entidade geográfica, através de três formas básicas: pontos, linhas e polígonos, definidas por suas coordenadas cartesianas.

A partir das formas básicas definidas na seção 2.1.3, pode-se também formar outras geometrias como, polígonos complexos, que permitem buracos ou partes disjuntas; poliedro, limitado por quatro ou mais polígonos, denominados faces, sendo que as interseções das faces formam as arestas e as interseções das arestas formam os vértices.

A fronteira do polígono divide o plano em duas regiões: o interior e o exterior. Os polígonos são usados para representar unidades de dados geográficos espaciais individuais, como, setores censitários, distritos, zonas de endereçamento postal. A Figura 2.5 mostra as formas básicas utilizadas e as derivadas delas.

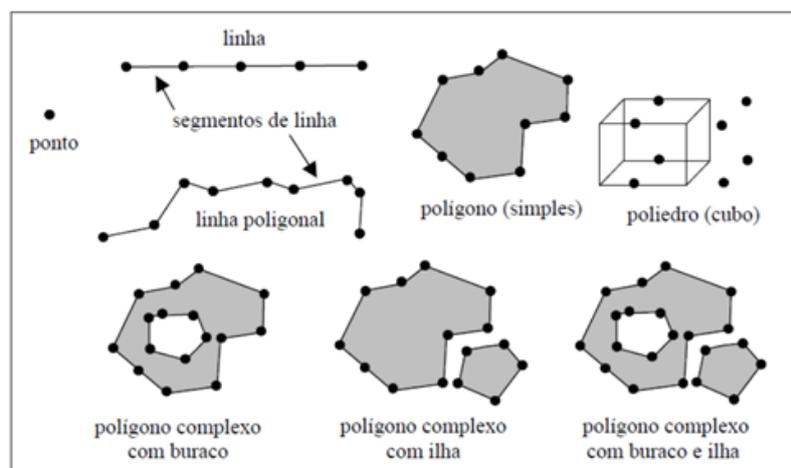


Figura 2.5: Formas Derivados dos Dados Espaciais a partir dos Básicos [1].

Nesse trabalho, são tratados os dados geográficos, os quais são representados de forma vetorial; onde os dados geométricos em sistemas de banco de dados fornece uma abstração fundamental para modelar a estrutura geométrica de objetos no espaço, suas relações, propriedades e operações, utilizadas para representar as coordenadas das fronteiras de cada entidade ou objeto geográfico, através de três formas básicas apresentadas acima.

2.3 Banco de Dados Geográficos

Um banco de dados geográficos, segundo Hartmur [13], é um sistema de base de dados, que oferece tipos de dados espaciais em seu modelo de dados e sua linguagem de consulta suporta tipos de dados espaciais em sua implementação, fornecendo pelo menos indexação espacial e algoritmos eficientes para integração espacial. Resumidamente, um banco de dados geográficos é aquele que tem a capacidade de armazenar dados geográficos, além dos dados convencionais ou alfanuméricos.

Um banco de dados geográfico pode estar organizado de diferentes formas, desde um banco de dados constituído de um conjunto de arquivos até sua implementação em um Sistema Gerenciador de Banco de Dados (SGBD)¹. Desta forma, em relação ao armazenamento dos dados geográficos, é possível ter dados geográficos em arquivo e dados em SGBD, chamado de arquitetura dual [14], ou ter uma arquitetura integrada onde todos os dados são armazenados em um SGBD.

Essa dissertação trata especificamente dos bancos de dados geográficos organizados totalmente em SGBD. A vantagem dessa arquitetura é que o SGBD é produzido com características como controle de concorrência, recuperação a falhas, processamento de consulta com algoritmos específicos o que facilita a manutenção dos bancos de dados. Além dessas características, colocar todos os dados espaciais ou não em um SGBD tem a vantagem de permitir integração dos dados.

O primeiro passo na implementação de um banco de dados é a especificação de um projeto. A tarefa de projetar um banco de dados geográfico é árdua para a maioria das organizações públicas e privadas, em decorrência da complexidade das informações dessas instituições. Então é necessário que o projeto de banco de dados seja realizado com um modelo de dados de alto nível, conhecido como modelo conceitual.

2.4 Modelagens Conceituais de Dados Geográficos

Um modelo de dados é uma coleção de conceitos e operações utilizados para descrever e manipular um conjunto de dados. No projeto de banco de dados, são utilizados modelos de dados em diferentes níveis de abstração e etapas de desenvolvimento (projeto conceitual, lógico e físico), tendo modelos de alto nível, por exemplo, o modelo de Entidade-Relacionamento, até modelos mais específicos, relacionados ao armazenamento dos dados.

¹Um SGBD é um software que tem como objetivo auxiliar os desenvolvedores na criação e manutenção dos seus bancos de dados específicos. Dentre esses SGDBS destacam-se o Oracle, DB2, SQL Server, PostgreSQL, dentre outros.

No contexto da representação da geoinformação, é necessária a aplicação do paradigma dos quatro universos [1], apresentada na seção 2.1. Este paradigma procura descrever em seus quatro passos um objeto qualquer do mundo real na sua representação computacional.

A utilização do paradigma dos quatro universos como ferramenta de modelagem é uma forma de interpretar e compreender cenários reais e representá-los em meios computacionais.

2.4.1 Técnica de Modelagem de Objetos para Aplicações Geográficas

O modelo OMT-G (*Object Modeling Technique for Geographic Applications*) foi criado a partir do modelo OMT introduzindo primitivas geográficas com objetivo de aumentar a capacidade de representação semântica do espaço [15]. O modelo OMT-G provê primitivas para modelar a geometria e a topologia dos dados geográficos, oferecendo suporte às estruturas topológicas, estruturas de rede, múltiplas representações de objetos e relacionamentos espaciais. Além disso, o modelo permite a especificação de atributos alfanuméricos e métodos associados para cada classe.

As características principais do modelo OMT-G são: sua grande expressividade gráfica e sua capacidade de codificação, denotando a dinâmica da interação entre os objetos espaciais e não espaciais. O modelo é baseado em três conceitos principais: classes, relacionamentos e restrições de integridade espaciais. Classes e relacionamentos definem as primitivas básicas usadas para criar esquemas estáticos da aplicação. Além disso, OMT-G propõe o uso de três diferentes diagramas no processo de desenvolvimento de uma aplicação geográfica: diagrama de classes, diagrama de transformação e diagrama de apresentação. Como um dos objetivos desse trabalho é a definição do projeto de banco de dados, o diagrama de classes será utilizado[2].

É muito importante no projeto de uma aplicação que as restrições de integridade espacial sejam identificadas, garantindo as condições que são necessárias para a integridade do banco de dados. Podem ser citados, como principais tipos de restrições de integridade, que constantemente estão presentes na modelagem de bancos de dados convencionais, as restrições de domínio, de chave, de integridade referencial e de integridade semântica. O objetivo dessa classificação pode ser visto como sendo o de abarcar as particularidades dos dados espaciais, onde inclui-se restrições topológicas, semânticas e definidas pelo usuário.

As restrições de integridade semânticas estão relacionadas ao significado implícito às feições geográficas, por exemplo, uma regra que impede que trechos de logradouro interceptem edifícios. Já as restrições de integridade definidas pelo usuário atuam como regras de negócio, permitindo assim manter a consistência do banco de dados. Como exemplo do uso desta restrição, pode-se citar a restrição da localização de postos de gasolina, que legalmente, necessitam estar a pelo menos 200 metros de distância de qualquer escola. Restrições definidas pelo usuário podem ser armazenadas e garantidas por um repositório ativo.

As primitivas do modelo OMT-G utilizadas para criar o diagrama de classes para aplicações geográficas, estão descritas a seguir.

Classes

As classes definidas pelo modelo OMT-G representam os três grandes grupos de dados (contínuos, discretos e não-espaciais) encontradas nas aplicações geográficas, oferecendo assim, uma visão integrada do espaço modelado. Suas classes podem ser georreferenciadas ou convencionais [2].

Essa diferenciação entre classes convencionais e georreferenciadas permite que diferentes aplicações compartilhem dados não espaciais, desta forma facilita o desenvolvimento de aplicações integradas e a reutilização de dados. A classe georreferenciada descreve um conjunto de objetos que possuem representação espacial e estão associados a regiões da superfície da terra[16], representando a visão de campos e de objetos. A classe convencional descreve um conjunto de objetos com propriedades, comportamento, relacionamentos, e semântica semelhantes, e que possuem alguma relação com os objetos espaciais, mas que não possuem propriedades geométricas [2].

A diferença entre as classes convencionais e georreferenciadas é apresentada na Figura 2.6. Pode-se observar que a classe georreferenciada tem um quadrado no lado superior esquerdo onde é representado o tipo de dado geométrico.

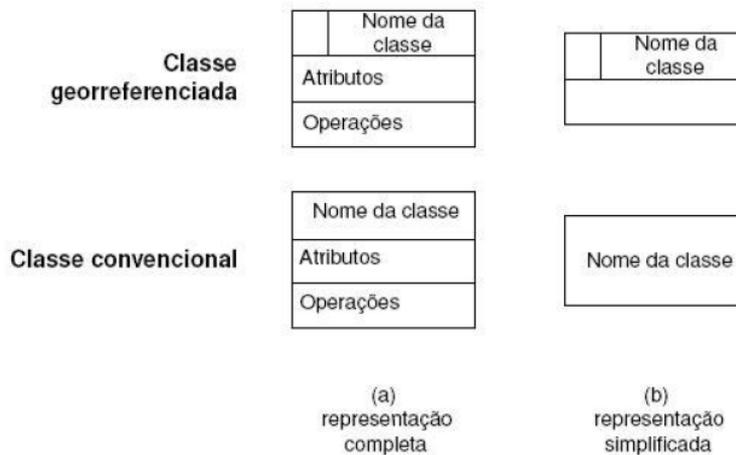


Figura 2.6: Notação Gráfica para Classes no modelo OMT-G [2].

O modelo OMT-G apresenta um conjunto fixo de alternativas de representação geométrica, usando uma simbologia que distingue geo-objetos e geo-campos. O modelo OMT-G define cinco classes descendentes de geo-campo: isolinhas, subdivisão planar, tesselação, amostragem e malha triangular (*triangulated irregular network*, TIN) apresentada na Figura 2.7, e duas classes descendentes de geo-objeto: geo-objeto com geometria e geo-objeto com geometria e topologia [2].

A classe geo-objeto com geometria representa objetos que possuem apenas propriedades geométricas, e é especializada em classes: Ponto, Linha e Polígono. Como exemplo citamos, respectivamente, árvore, meio-fio e curso de água apresentada na Figura 2.8 (a).

A classe geo-objeto com geometria e topologia representa objetos que possuem, além das propriedades geométricas, propriedades de conectividade topológica, sendo especificamente voltadas para a representação de estruturas em rede, tais como sistemas de abastecimento de água ou fornecimento de energia elétrica. Essas propriedades estão

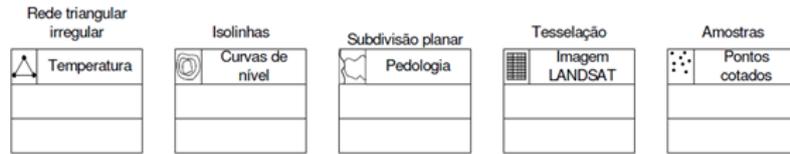


Figura 2.7: Geo-campos [2].

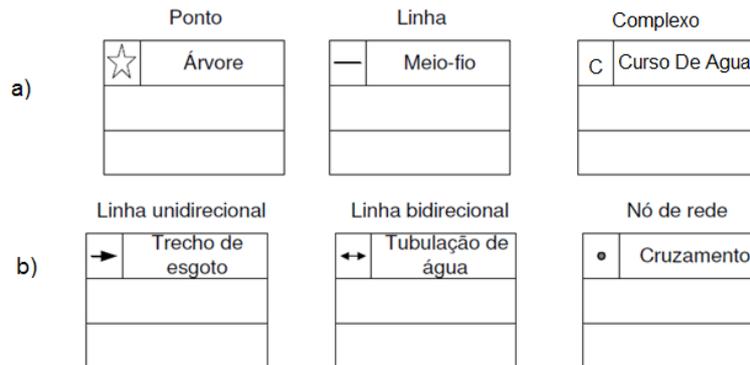


Figura 2.8: Geo-objetos (a) com geometria, (b) com geometria e topologia[2].

presentes em classes descendentes que representam nós e arcos, da forma usualmente adotada na teoria dos grafos. Os arcos podem ser unidirecionais, como em redes de esgoto, ou bidirecionais, como em redes de telecomunicações. Assim, as especializações previstas são denominadas nó de rede, arco unidirecional e arco bidirecional. Os segmentos orientados traduzem o sentido do fluxo da rede, se unidirecional ou bidirecional, dando mais semântica à representação [2], como exemplo apresentamos a Figura 2.8 (b).

Relacionamentos

Um problema na maioria dos modelos de dados existentes é o fato deles ignorarem a possibilidade de modelagem dos relacionamentos entre fenômenos do mundo real. Considerando a importância das relações espaciais e não espaciais na compreensão do espaço modelado, o modelo OMT-G representa três tipos de relacionamentos entre suas classes: associações simples, relacionamentos topológicos em rede e relacionamentos espaciais. A distinção de tais relacionamentos objetiva definir explicitamente o tipo de interação que ocorre entre as classes [2].

Associações simples representam relacionamentos estruturais entre objetos de classes diferentes, convencionais ou georreferenciadas. Relacionamentos espaciais representam relações topológicas, métricas, de ordem e *fuzzy*. Algumas relações podem ser derivadas automaticamente, a partir da forma geométrica do objeto, no momento da entrada de dados ou da execução de alguma análise espacial, como o relacionamento topológico. As relações explícitas, no entanto, precisam ser expressamente especificadas pelo usuário, de modo a permitir que o sistema armazene e mantenha atualizada aquela informação [2].

No modelo OMT-G, associações simples são indicadas por linhas contínuas, enquanto relacionamentos espaciais são indicados por linhas pontilhadas a Figura 2.9 a/b mostra

exemplos. Isso torna fácil a distinção visual entre relacionamentos baseados em atributos alfanuméricos e baseados na localização e forma geométrica dos objetos. O nome do relacionamento é anotado sobre a linha, e uma seta usada para deixar clara a direção de leitura, por exemplo, na Figura 2.9 b, lê-se “lote contém edificação”.

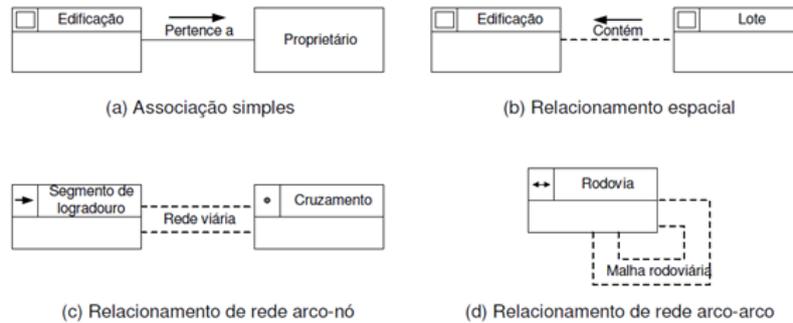


Figura 2.9: Relacionamentos [1].

Os relacionamentos de rede são relacionamentos entre objetos que estão conectados uns com os outros. Relacionamentos de rede são indicados por duas linhas pontilhadas paralelas, entre as quais o nome do relacionamento é explicitado (Figura 2.9 c). Os relacionamentos são em geral especificados entre uma classe de nós e uma classe de arcos, mas estruturas de redes sem nós podem ser definidas, especificando um relacionamento recursivo sobre uma classe de arcos (Figura 2.9 d) [2].

O modelo OMT-G considera um conjunto de relacionamentos espaciais entre classes georreferenciadas, identificando um conjunto mínimo de relacionamentos espaciais, que compreende somente cinco relacionamentos, a partir dos quais todos os outros podem ser especificados: toca, em, cruza, sobrepõe e disjunto. Eventualmente, é necessário aplicar um conjunto maior de relacionamentos devido a fatores culturais ou semânticos que são familiares para os usuários, incluindo relacionamentos de significado difuso, tais como perto de, ou ao norte de [2].

Alguns relacionamentos dependem da representação geométrica e assim, só existem entre determinadas classes. Por exemplo, o relacionamento contém pressupõe que uma das classes envolvidas seja um polígono. Neste ponto, as aplicações tradicionais diferem das geográficas, em que não existem esses tipos de restrições para as associações entre classes convencionais, sendo independente de fatores como comportamento geométrico [2].

O conjunto de conceitos que o usuário tem sobre cada objeto do mundo real sugere uma determinada representação porque existe uma interdependência entre a representação, o tipo de interpretação e a finalidade que será dada a cada entidade geográfica. No modelo OMT-G, isto é considerado para que sejam estabelecidas as relações que envolvem classes georreferenciadas [2].

As representações de cardinalidade, generalização e especificação em OMT-G seguem o mesmo padrão da diagramação UML.

2.4.2 Infraestrutura Nacional de Dados Espaciais - INDE

Essa seção apresenta uma visão geral da Infraestrutura de Dados Espaciais (IDE) a fim de contextualizar os modelos de dados especificados nessa dissertação dentro das

recomendações da INDE, que é a Infraestrutura Nacional de Dados Espaciais. Nesse sentido, para definir uma IDE são utilizados, alguns significados especificados a seguir.

O Comitê Federal dos Estados Unidos definiu sua Infraestrutura Nacional de Dados Espaciais ou *National Spatial Data Infrastructure* (NSDI) como “um conjunto de políticas, padrões e procedimentos sob os quais organizações e tecnologias interagem para promover o uso, administração e produção mais eficientes de dados geoespaciais” [17].

O Conselho de Informação Espacial da Austrália e Nova Zelândia (ANZLIC) responsável pela implantação da IDE de Australiana destaca que: “Uma Infraestrutura de Dados Espaciais provê uma base para busca de dados espaciais, avaliação, transferência e aplicação para os usuários e provedores dentro de todos os níveis de governo, do setor comercial e industrial, dos setores não lucrativos, acadêmicos e do público geral” [18].

No Brasil, a definição de IDE adotada é: “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” [1].

Considerando pequenas diferenças nas definições acima, elas entram em consenso ao reportarem que uma IDE deve tratar com um conjunto de tecnologias para disseminação da informação em setores governamentais, não governamentais, industriais, não lucrativos e acadêmicos.

No Brasil, a INDE foi legalmente constituída pelo Decreto Presidencial No 6.666, de 27 de novembro de 2008, com o objetivo de formular definições, apontar responsabilidades, e estabelecer diretrizes para o desenvolvimento de aplicativos com dados espaciais.

A Comissão Nacional de Cartografia (CONCAR) votou pela criação de um Comitê Técnico responsável pela elaboração do plano de ação para a implantação da INDE. Este comitê denominado Comitê para o planejamento da INDE (CINDE) foi constituído entre Janeiro e Março de 2009. Tendo como produto o Plano de Ação Para Implantação da Infraestrutura Nacional de Dados Espaciais, documento oficial que formula definições, aponta responsabilidades e especifica diretrizes para uma implantação de uma IDE no Brasil [6].

Segundo o comitê [6], uma IDE deve ser desenvolvida com o objetivo de atingir os seguintes objetivos:

1. Compartilhar informação geográfica, inicialmente na administração pública e depois para toda a sociedade.
2. Incrementar a administração eletrônica no setor público.
3. Incorporar a informação geográfica produzida pela iniciativa privada.
4. Harmonizar a informação geográfica, bem como registrar as características dessa informação geográfica.
5. Subsidiar a tomada de decisões de forma mais eficiente e eficaz.

A justificativa para a implantação de uma IDE está ligada, fundamentalmente, a duas ideias [19]:

1. O acesso aos dados geográficos existentes deve ocorrer de modo fácil, cômodo e eficaz.

2. A informação geográfica deve ser reutilizada uma vez que tenha sido usada para o projeto que justificou a sua aquisição, face aos custos elevados de sua produção.

É consenso internacional que uma IDE deve estar fundamentada em cinco componentes, os quais são fortemente relacionados e interagem entre si. A Figura 2.10 apresenta esses componentes e serviu de base para a elaboração do Plano de Ação [6].



Figura 2.10: Pilares de uma IDE [1].

Os dados são a parte central de toda IDE. Compreendem vários conjuntos de dados geoespaciais dos quais produzem informações.

As pessoas, são as partes envolvidas ou interessadas na produção de uma IDE, são também chamados atores. O setor público e o setor privado são responsáveis pela aquisição, produção, manutenção e oferta de dados espaciais. O setor acadêmico é o responsável pela educação, capacitação, treinamento e pesquisa em IDE. Por último, o usuário determina que dados espaciais são requeridos e como devem ser acessados.

O componente institucional compreende as questões de política, legislação e coordenação. Essas questões são importantes para a padronização dos dados de acordo com a especificação da IDE brasileira.

Além disso, a tecnologia descreve as ferramentas utilizadas por uma IDE tanto quanto a arquitetura que ela deve ter para alcançar os objetivos propostos de acesso distribuição e armazenamento de dados espaciais.

Por ultimo, o documento das Normas e Padrões permite a normatização e padronização de estruturas de dados espaciais para facilitar o intercâmbio de informações entre IDE's regidas pelo Plano de Ação brasileiro e talvez com outras IDE's pelo mundo [6].

2.4.3 Documento das Especificações Técnicas para a Estruturação de Dados Geospaciais e Vetoriais

O documento das especificações na sua versão 2.0 tem como objetivo padronizar estruturas de dados que viabilizem o compartilhamento de dados, a interoperabilidade e a racionalização de recursos entre os produtores e usuários de dados e informação cartográfica.

Segundo as especificações da INDE a estrutura de dados geospaciais vetoriais foram modeladas com a técnica orientada a objetos [3]. Assim todo dado geoespacial recebe a denominação de classe de objetos. Na modelagem conceitual dos dados geospaciais, as ocorrências são representadas por classes de objetos de mesma natureza e funcionalidade. Estas classes foram agrupadas em categorias de informação, cuja premissa básica é o aspecto funcional comum [3]. As categorias de informação são apresentadas na Figura 2.11.

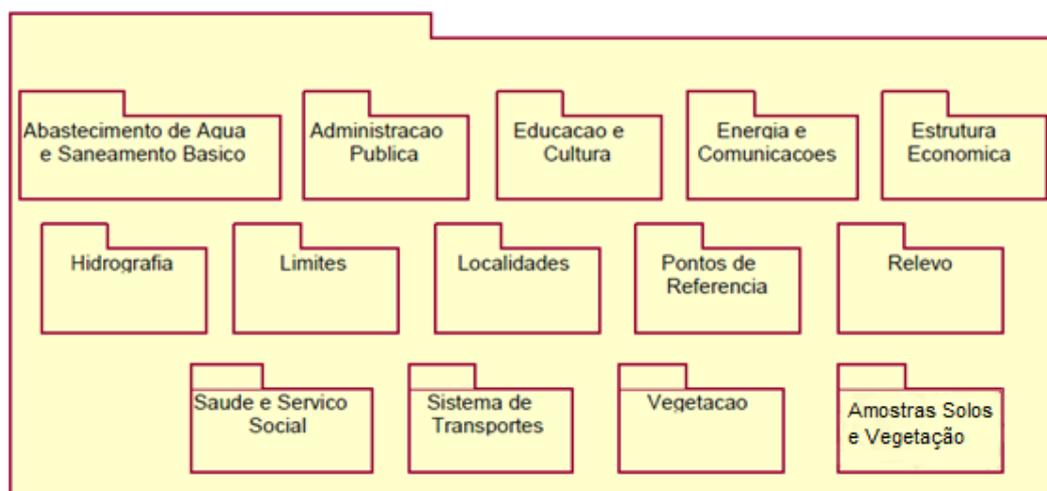


Figura 2.11: Categorias de informação da CONCAR [3].

- Hidrografia: Categoria que representa o conjunto das águas interiores e oceânicas da superfície terrestre, bem como elementos, naturais ou artificiais, emersos ou submersos, contidos nesse ambiente.
- Relevo: Categoria que representa a forma da superfície da Terra e do fundo das águas tratando, também, os materiais expostos, com exceção da cobertura vegetal.
- Vegetação: Categoria que representa, em caráter geral, os diversos tipos de vegetação natural e cultivada.
- Sistema de Transportes: Categoria que agrupa o conjunto de sistemas destinados ao transporte e deslocamento de carga e passageiros, bem como as estruturas de suporte ligadas a estas atividades.
- Energia e Comunicações: Categoria que representa as estruturas associadas à geração, transmissão e distribuição de energia, bem como as de comunicação.

- **Abastecimento de Água e Saneamento Básico:** Categoria que agrupa o conjunto de estruturas associadas à captação, ao armazenamento, ao tratamento e à distribuição de água, bem como as relativas ao saneamento básico.
- **Educação e Cultura:** Categoria que representa as áreas e as edificações associadas à educação e ao esporte, à cultura e ao lazer.
- **Estrutura Económica:** Categoria que representa as áreas e as edificações onde são realizadas atividades para produção de bens e serviços que, em geral, apresentam resultado econômico.
- **Localidades:** Categoria que representa os diversos tipos de concentração de habitações humanas.
- **Pontos de Referência:** Categoria que agrupa as classes de elementos que servem como referência à medições em relação à superfície da Terra ou de fenômenos naturais.
- **Limites:** Categoria que representa os distintos níveis político administrativos e as áreas especiais; áreas de planejamento operacional, áreas particulares (não classificadas nas demais categorias), bem como os elementos que delimitam materialmente estas linhas no terreno.
- **Administração Pública:** Categoria que representa as áreas e as edificações onde são realizadas as atividades inerentes ao poder público.
- **Saúde e Serviço Social:** Categoria que representa as áreas e as edificações relativas ao serviço social e à saúde.

2.4.4 *Open Geospatial Consortium (OGC)*

A OGC é uma organização voluntária internacional de padrões de consenso. No OGC, mais de 400 organizações comerciais, governamentais, não lucrativas e instituições de pesquisa do mundo todo colaboram num processo de consenso aberto encorajando o desenvolvimento e a implementação de padrões para conteúdo e serviços geomáticos, SIG, processamento de dados e troca [20]. Os padrões estabelecidos pela OGC suportam soluções interoperáveis baseadas na internet e que também são desenvolvidos para trabalhar com informações espaciais complexas e serviços utilizados por diferentes tipos de aplicações [11].

Os padrões atualmente aprovados pela OGC estão agrupados em conjuntos temáticos de serviços de catálogo, serviços de processamento, codificação, serviços de dados, serviços de retrato (imagem) e outros padrões comuns. Em seguida, são descritos os mais importantes:

- *Geographic Markup Language (GML):* Um padrão de codificação baseado em *eXtensible Markup Language (XML)* para informações geográficas desenvolvidas pelo consórcio OGC. A GML foi especificada para o transporte e armazenamento de informação geográfica, incluindo propriedades espaciais e não espaciais das feições geográficas. O objetivo da GML é oferecer um conjunto de regras com as quais um usuário pode definir seu esquema para descrever seus dados. Sua versão 3.0 inclui esquemas que contêm os modelos de geometria, feições e superfícies.

- *Keyhole Markup Language* (KML): uma notação XML para exibir dados geográficos em um navegador da Terra, como Google Earth, Google Maps. O KML utiliza uma estrutura de marcadores com elementos e atributos aninhados e se baseia no padrão XML. KML foi desenvolvido para uso com o Google Earth, sendo originalmente chamado *Keyhole Earth Viewer*. Foi criado por Keyhole Inc. Sendo um padrão internacional da OGC, o arquivo KML especifica um conjunto de características (marca local, imagens, polígonos, modelos 3D, descrições textuais, entre outros) para exibição no Google Earth, Maps e Mobile. KML compartilha algumas das mesmas gramáticas estruturais como GML.
- *Web Map Service*(WMS): é um padrão que fornece uma interface HTTP simples para solicitar imagens de mapas georreferenciadas de um ou mais bancos de dados geoespaciais distribuídos. A solicitação WMS define as camadas geográfica(s) e área de interesse a ser processado. A resposta ao pedido é uma ou mais imagens georreferenciadas do mapa (retornado como JPEG, PNG) que podem ser exibidos num navegador. A interface também suporta a capacidade para especificar se as imagens retornadas devem ser transparentes ou não para que as camadas de vários servidores possam ser combinadas.
- *Web Feature Services*(WFS): é um serviço padrão, que fornece uma interface de comunicação que permite interagir com mapas atendidas pelo padrão WMS, por exemplo, editar a imagem oferecida pelo serviço WMS ou analisar a imagem usando critérios geográficos. Para realizar essas operações, utiliza-se GML linguagem derivada do XML, que é o padrão pelo qual são transmitidas as ordens WFS, no entanto, qualquer outro formato vetorial pode ser utilizado.

WFS não transacional permite a consulta e recuperação de recursos geográficos. Pelo contrário, WFS-T (*Web Feature Service Transacional*) permite também criar, apagar e atualizar esses recursos de mapas geográficos.

- *Web Coverage Service*(WCS): define uma interface padrão e operações que permitem o acesso interoperável para dados geoespaciais do tipo “matriz”. O termo “matriz” normalmente refere-se a conteúdos como imagens de satélite, fotos aéreas digitais e dados de elevação digital. Este padrão também possui a definição de operação de transação, onde opcionalmente pode ser implementado por servidores WCS. Esta operação de transação permite ao usuário adicional, modificar e eliminar dados matriciais num servidor WCS. As referências de transação ou operação de solicitação incluem a criação de novo dado matricial ou modificações de um existente, abrangendo todos os metadados necessários para um dado matricial.

2.5 Arquiteturas Orientadas a Serviços

Segundo Bieberstein [21], para dar uma definição sobre uma Arquitetura Orientada a Serviços (SOA), do inglês *Service Oriented Architecture*, são precisos alguns conceitos básicos. Primeiro, um serviço é uma manifestação de uma tarefa de negócios. E quando é realizado por meio das tecnologias, o termo serviço aplica-se a um recurso de software com uma especificação serviço.

Um serviço está disponível para consulta, ligação e invocação por um consumidor de serviços. O prestador de serviço realiza a implementação da especificação do serviço. Segundo, com base na definição de serviço, orientação a serviços é uma forma de integrar um negócio como um conjunto de serviços ligados. Então SOA é um estilo de arquitetura que suporta a orientação a serviços. Estes serviços comunicam-se uns com os outros. A comunicação pode envolver a passagem de dados simples ou pode envolver dois ou mais serviços de coordenação de alguma atividade. Por último, um aplicativo é composto por um conjunto de serviços integrados.

Em e-Ping [7], SOA é definido como uma arquitetura para interoperabilidade de sistemas por meio de conjunto de interfaces de serviços fracamente acoplados, onde os serviços não necessitam de detalhes técnicos da plataforma dos outros serviços para a troca de informações ser realizada. A arquitetura SOA permite a criação de sistemas de informação altamente escaláveis, que refletem o negócio da organização, por sua vez fornece uma forma bem definida de exposição e invocação de serviços, o qual facilita a interação entre diferentes sistemas.

De forma geral, Bieberstein [21] define as seguintes camadas de software numa SOA:

- Aplicações básicas: Sistemas desenvolvidos em qualquer arquitetura ou tecnologia, geograficamente dispersos.
- Exposição de funcionalidade: Onde as funções da camada aplicada são expostas como um serviço (normalmente como *web services*).
- Integração de serviços: Facilitam o intercâmbio de dados entre diferentes serviços.
- Processo de composição: O processo de composição, o qual define o processo em termos de negócio e suas necessidades, e que varia de acordo com o negócio.
- Entrega: Camada onde os serviços são implantados para usuários finais.

A arquitetura SOA é tanto um marco de trabalho para o desenvolvimento de software, como de implementação. Para que um projeto SOA tenha sucesso, os desenvolvedores de software devem criar serviços comuns, que são orquestrados por clientes ou middleware para implementar processos de negócios.

Uma adoção de arquitetura SOA tem benefícios como melhoria no tempo de fazer mudanças nos processos, facilidade para evoluir aos modelos de negócios baseados em *outsourcing*², poder de substituir elementos da camada de aplicação SOA sem interrupção do processo de negócio e facilidade para a integração de tecnologias diferentes [21].

SOA para Integração de Dados Espaciais

Proposta por Sha Zongyao e Xie Yichun em [4], *Service-Oriented Architecture for Spatial Data Integration* (SOA-SDI) é uma arquitetura para integração de dados espaciais no contexto de um grande número de fontes disponíveis de dados espaciais, que estão fisicamente armazenados em lugares distintos, em formatos e plataformas diferentes.

A arquitetura SOA-SDI conecta e coordena múltiplas fontes de dados espaciais, integra serviços de ferramentas de processamento de dados, analisa e apresenta os dados em um

²Refere-se ao processo de contratação a terceiros

formato desejado através de Sistemas de Informação Geográfica (SIG) sobre a Internet, Intranet ou *Wide Area Network*(WAN).

A arquitetura SOA-SDI mostrada na Figura 2.12 propõe quatro camadas ou componentes, que são: camada de múltiplos provedores de serviços, camada de integração de dados, camada de serviços de infraestrutura (*Backend Services*) e a *interface front-end* para a representação de dados espaciais.

A primeira camada está representada por múltiplos provedores de serviços, cada provedor de serviços oferece um conjunto de interfaces através da qual as aplicações clientes podem extrair dados remotos e manipulá-los sobre a rede. A arquitetura SOA-SDI considera esta camada como a componente chave, pela grande quantidade de dados geográficos em diferentes organizações que tem sido acumulado.

Em SOA-SDI, as aplicações clientes acessam os dados geográficos através de um conjunto de *web services* através dos provedores de dados. Como apresentado anteriormente, o OGC propõem várias especificações para compartilhar dados espaciais e sua interoperabilidade baseado em estratégias orientadas a serviços. Entre eles, tem WMS e WFS, que foram amplamente aplicados para acessar dados espaciais em diferentes fontes sobre a internet.

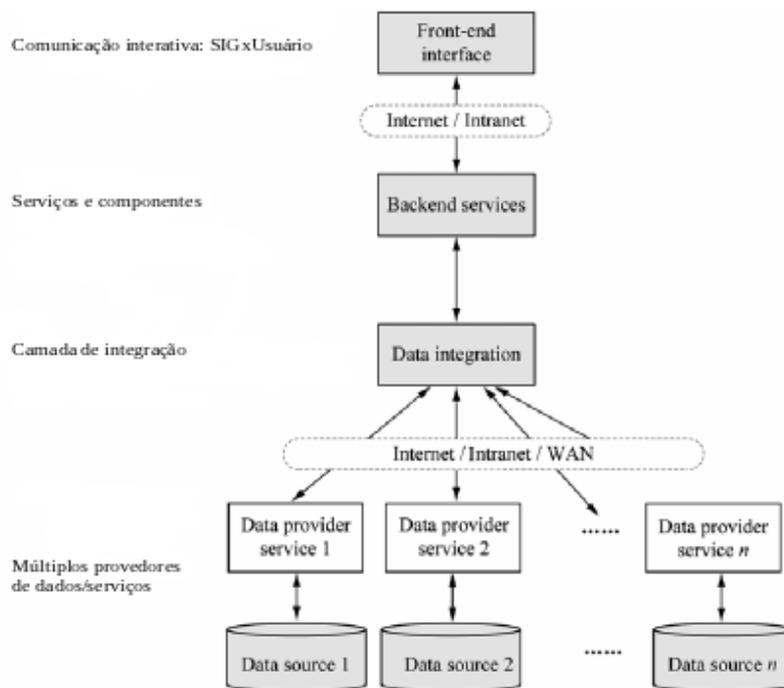


Figura 2.12: Arquitetura SOA-SDI de Quatro Camadas [4].

Geralmente o processo de consumo envolve dois passos: pedido e resposta. O primeiro passo é fazer um pedido das aplicações clientes ao provedor de serviços com o formato XML. A resposta também tem o formato XML e provém do provedor de serviços para a aplicação cliente.

A segunda camada, de integração de dados, coordena múltiplos provedores de serviços de dados para diferentes aplicações. A principal função desse componente é a construção de pedidos de dados em um formato apropriado que será enviado através da rede de

comunicação de dados para serviços específicos de dados do provedor, e analisa a resposta retornada desses serviços. Para algum pedido de uma aplicação cliente, a camada de integração de dados é responsável por enviar um pedido bem formatado que o provedor de serviços de dados entenda. Para cada resposta, a camada de integração de dados é responsável de extrair os dados interessados para a camada superior.

A camada *Backend services* é composta por uma série de serviços e componentes na qual as aplicações SIG estão construídas. Os componentes podem ser dinamicamente atualizados, removidos ou aderidos.

Por último, a camada de *interface front-end* prove uma comunicação interativa entre os sistemas SIG e os usuários finais. É composta por elementos GUI, por exemplo, botões, formulários, mapas de imagens.

VIRTUAL DATABASE

Esta arquitetura foi apresentada por Marcel Frehner e Martin Brandi [5] no ano 2006. O foco principal deste trabalho foi a extensão das técnicas de recuperação, análise, e visualização de dados ambientais espacialmente relacionados com base na integração de repositórios de dados ecológicos, oferecendo capacidades de análise espacial aplicável aos conjuntos de dados distribuídos.

Sua implementação está baseada em módulos e seu projeto é baseado em arquitetura orientada a serviço.

A Figura 2.13 mostra a sua arquitetura em três níveis: nível de administração de recursos, nível de aplicação e nível de apresentação. Cada uma dessas camadas é descrita a seguir.

No nível de administração de recursos estão os repositórios de dados distribuídos com diferentes tipos de dados e sistemas de armazenamento. Os repositórios estão armazenados em diferentes instituições da Suíça. Acima de cada repositório existe uma camada de acesso, que manipula os dados de acordo com o banco de dados. A camada de acesso oculta a heterogeneidade dos dados através de interfaces uniformes e padronizadas que são acessadas mediante a camada de integração.

O nível de aplicação é responsável pela integração de dados. A camada de integração reúne e integra os dados recuperados a partir das camadas de acesso, a fim de tornar a distribuição transparente de dados. Incluído na camada de integração encontra-se um repositório persistente armazenando informações de manutenção e atuando como um cachê de dados para recuperação rápida de dados. A lógica de negócios opera sobre o conjunto combinado de dados ignorando origens possivelmente diferentes. A camada de aplicação deve oferecer componentes de negócios para a recuperação, consulta, análise e visualização de dados ecológicos.

A finalidade principal de componentes relacionados com a exibição é para compilar e processar mapas cartográficos mostrando fenômenos ecológicos básicos e características topográficas. Estes são métodos relacionados a sobreposições espaciais combinando com capacidades *buffering*. Sobreposições espaciais permitem que se cruzem mutuamente dados ecológicos disponíveis. Os componentes de negócios são acessados pelos serviços correspondentes dentro da camada de aplicação (catálogo, análise espacial, serviço de mapas). A aplicação de web services oferece uma solução a este respeito.

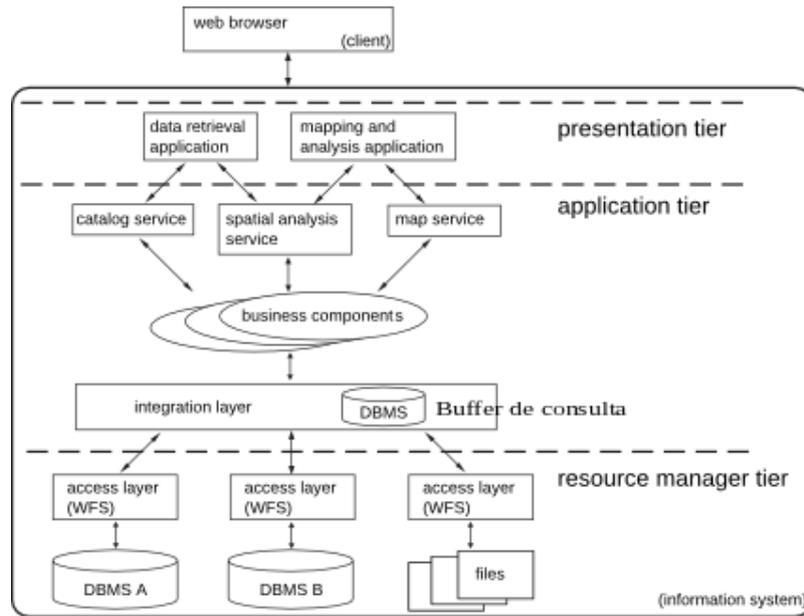


Figura 2.13: Arquitetura Orientada a Serviços Baseada em Três Níveis [5].

Uma vez que muitos programas, independente da plataforma, são capazes de construir e interpretar as mensagens *Simple Object Access Protocol* (SOAP), a arquitetura do banco de dados virtual permite o acesso a componentes de negócios.

Por último a camada de apresentação contém aplicativos que lidam com a interação do usuário e exibição. Estas aplicações web usam serviços expostos pela camada de aplicação. Tem dois módulos de aplicações ligadas: uma ferramenta que oferece suporte à recuperação de dados e outro para o mapeamento e análise.

Arquitetura do Plano de Ação da Infraestrutura Nacional de Dados Espaciais

Como resultado do trabalho do Comitê para planejamento da Infraestrutura Nacional de Dados Espaciais (CINDE), foi instituído o plano de ação para implantação da INDE com o fim de ordenar a geração, o armazenamento, o acesso, a disseminação, o compartilhamento e o uso dos dados geoespaciais de origem federal, estadual, distrital e municipal.

Neste plano de ação [6] apresentam uma arquitetura SOA que utilizam os padrões OGC, conforme ilustrado na Figura 2.14.

O primeiro nível abrange fontes de dados espaciais localizadas em distintos lugares, elas podem pertencer a universidades, Organização não Governamental (ONG), governo federal, governo municipal ou provedores comerciais de dados.

No segundo nível, serviços web OGC, o acesso aos dados é realizado apenas através de serviços. Nessa camada o usuário não precisaria conhecer o local onde os dados estão armazenados, pois cada provedor de dados se encarrega de registrar, junto a um serviço de catálogo, que dados possui, onde estão, como estão organizados e onde estão os metadados.

O último nível é composto pelos usuários e o catálogo de serviços (registro de serviços OGC) onde os usuários obtêm as primeiras informações sobre os serviços que estão pro-

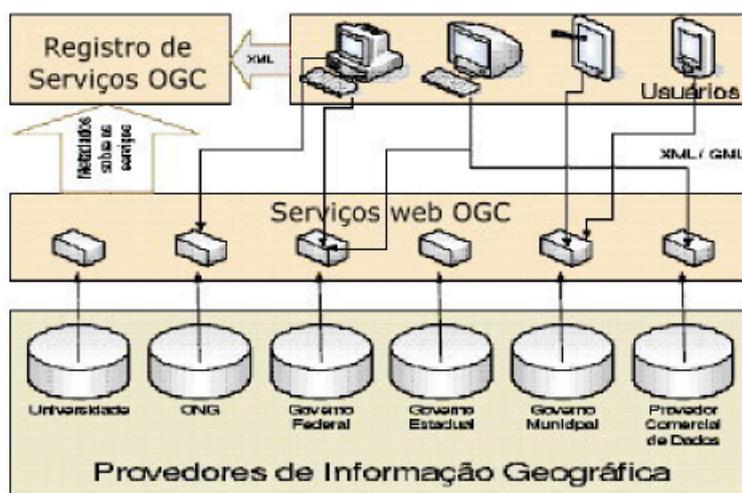


Figura 2.14: Arquitetura Proposta no Plano de Ação da INDE [6].

curando. A ideia geral das três arquiteturas é oferecer interoperabilidade e a integração de dados espaciais de fontes de dados heterogêneas e implementadas em distintos SGBD.

2.6 Sistema de Informação Geográfico para Web

Um Sistema de Informação Geográfico (SIG) é um sistema capaz de coletar, armazenar, manipular, recuperar e analisar dados espaciais [22] e conforme a Infraestrutura Nacional de Pesquisas Espaciais (INPE) é composta por ferramentas computacionais de geoprocessamento, que permitem a realização de análises complexas, ao integrar dados de diversas fontes e ao criar bancos de dados georreferenciadas. A diferença básica entre um sistema de informação convencional e um sistema de informação geográfico é que o SIG tem a capacidade de armazenar atributos espaciais, além dos atributos convencionais. Essa característica adicional dos SIG's permite ilustrar relações, conexões e padrões que não são facilmente visíveis em qualquer conjunto de dados, essa característica torna mais fácil a tomada de decisões nas organizações.

2.6.1 Arquitetura de um SIG

A Figura 2.15 mostra uma arquitetura de SIG [1]. Como pode ser observado, o banco de dados geográfico no nível mais baixo é administrado por um SGBDG oferecendo armazenamento e recuperação de dados espaciais e seus atributos, outros dados não espaciais são também armazenados da forma convencional. O SGBDG também oferece suporte às consultas espaciais, tendo funções espaciais para tratar os diferentes tipos de dados armazenados no banco de dados, por exemplo, funções que determinam se uma cidade limita com outra ou se um ponto pertence a uma determinada área. No nível intermediário a arquitetura mostra mecanismos de processamento de dados espaciais, como: a entrada e integração de dados, consulta e análise espacial. Por exemplo, para uma aplicação web os dados contidos no banco de dados devem ser convertidos para que a interação

do usuário com o mapa que representa os dados espaciais seja feito da forma mais natural possível, os mecanismos de visualização e plotagem para a apresentação dos dados. A parte de consulta e análise espacial inclui as operações topológicas, álgebras de mapas, estatística espacial, modelagem numérica de terreno, geodesia e fotogrametria, produção cartográfica, processamento de imagens e modelagem de redes. Dependendo das funcionalidades e requisitos dos sistemas, sua finalidade será mais específica para alguns casos e menos para outros [16]. Finalmente no nível intermediário, tem-se a parte de visualização e plotagem dos dados espaciais contidos no banco de dados. No nível mais próximo ao usuário encontra-se a interface com o usuário que define como o sistema é operado e controlado pelo usuário, nela ocorre toda a interação entre o usuário e a aplicação. Esta interface pode estar baseada na metáfora da “mesa de trabalho”, adaptada ao ambiente de navegação web, assim como baseada em linguagens de comando como *Spatial SQL* [1].

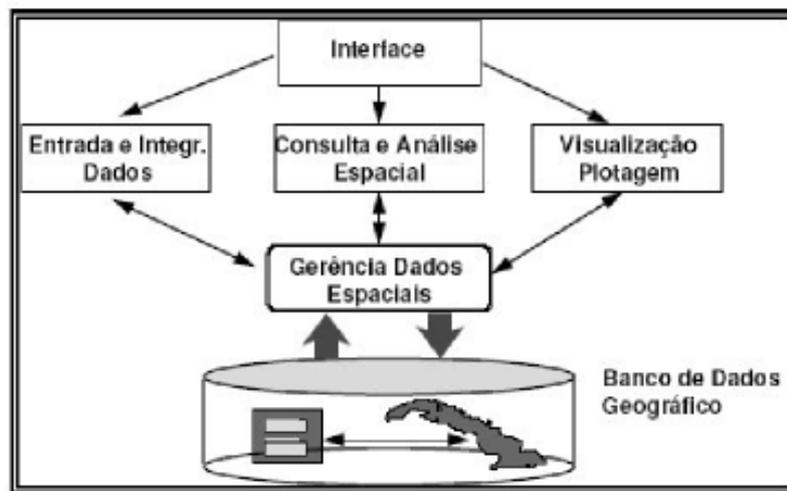


Figura 2.15: Arquitetura de Sistemas de Informações Geográficas [1].

2.6.2 Necessidades de Aplicações de Geoprocessamento

O setor de geoprocessamento no Brasil pode ser dividido em seis segmentos. Cada segmento representa uma área específica de atuação, em que cada grupo de funções de consulta apresentadas anteriormente se encaixa. Seguem as diferentes áreas em geoprocessamento [1]:

- Cadastral: aplicações de cadastro urbano e rural, realizadas tipicamente por Prefeituras, em escalas que usualmente variam de 1:1.000 a 1:20.000. A capacidade básica de SIG's para atender este setor é dispor de funções de consulta a bancos de dados espaciais e apresentação de mapas e imagens.
- Cartografia Automatizada: realizada por instituições produtoras de mapeamento básico e temático. Neste caso, é essencial dispor de ferramentas de aerofotogrametria digital e técnicas sofisticadas de entrada de dados (como digitalizadores ópticos) e de produção de mapas (como gravadores de filme de alta resolução).

- Ambiental: instituições ligadas às áreas de Agricultura, Meio-Ambiente, Ecologia e Planejamento Regional, que lidam com escalas típicas de 1:10.000 a 1:500.000. As capacidades básicas do SIG's para atender a este segmento são: integração de dados, gerenciamento e conversão entre projeções cartográficas, modelagem numérica de terreno, processamento de imagens e geração de cartas.
- Concessionárias/redes: Neste segmento, temos as concessionárias de serviços (Água, Energia Elétrica, Telefonia). As escalas de trabalho típicas variam entre 1:1.000 a 1:5.000. Cada aplicação de rede tem características próprias e com alta dependência de cada usuário. Os SIG's para redes devem apresentar duas características básicas: a forte ligação com bancos de dados relacionais e a capacidade de adaptação e personalização. O pacote básico disponível com os SIG's deste segmento é insuficiente para a realização da maioria das aplicações, pois cada usuário tem necessidades completamente distintas. Assim, os usuários deste setor realizam significativos desenvolvimentos nas linguagens de aplicação do SIG escolhido.
- Planejamento rural: Neste segmento, têm-se as empresas agropecuárias que necessitam planejar a produção e distribuição de seus produtos. As escalas de trabalho típicas variam entre 1:1.000 a 1:50.000. Cada aplicação tem características próprias e com alta dependência de cada usuário. Os SIG's devem apresentar duas características básicas: a forte ligação com bancos de dados relacionais e a capacidade de adaptação. O pacote básico disponível com os SIG's deste segmento é insuficiente para a realização da maioria das aplicações, pois cada usuário tem necessidades completamente distintas. Assim, os usuários deste setor realizam significativos desenvolvimentos nas linguagens de aplicação do SIG escolhido.
- *Business geographic*: Neste segmento, temos as empresas que necessitam distribuir equipes de vendas e promoção ou localizar novos nichos de mercado. As escalas de trabalho típicas variam entre 1:1.000 a 1:10.000. Cada aplicação tem características próprias e com alta dependência de cada usuário. As ferramentas de SIG devem prover meios de apresentação dos bancos de dados espaciais para fins de planejamento de negócios. Em especial, os SIG's devem ser adaptados ao cliente, com ferramentas de particionamento e segmentação do espaço para a localização de novos negócios e alocação de equipes.

Para o caso do projeto, pode-se notar que o segmento da área de processamento em que ele se encaixa é o ambiental. Assim, segundo as especificações do projeto, inicialmente as funcionalidades, demandaram a necessidade de utilizar apenas funções de consulta ao banco de dados georreferenciados e análise geográfica.

Capítulo 3

Modelos de dados e *Web Services*

Neste capítulo, são apresentadas as ferramentas e métodos utilizados durante a pesquisa, eles estão organizados nas seguintes seções: a seção 3.1 mostra a ferramenta utilizada para a modelagem dos dados geográficos. Na Seção 3.2 é visto o Sistema Gerenciador de Banco de Dados (SGBD). A Seção 3.3 apresenta o servidor de serviços geográficos que nos permitirá disponibilizar nossos dados geográficos. A Seção 3.4 mostra as extensões espaciais para alguns dos softwares utilizados durante a pesquisa. A Seção 3.5 mostra a ferramenta Quantum GIS, ferramenta que serviu nos testes iniciais. A Seção 3.6 apresenta uma ferramenta para extração, transformação e carregamento de dados desde distintas fontes de dados. Na Seção 3.7 é apresentado o Adobe Flex tecnologia, que suporta o desenvolvimento de aplicações ricas para a Internet. A Seção 3.8 é visto a ferramenta OpenScales que é um framework de mapas de código aberto (LGPL), permite construir aplicações ricas para internet (RIA) de mapas. Na Seção 3.9 é apresentado XGIS um software livre, construído sobre o framework de mapas OpenScales. Já na Seção 3.10 é apresentado a proposta da arquitetura orientada a serviços para o presente trabalho.

3.1 Ferramenta de Modelagem de Banco de Dados

A ferramenta de modelagem unificada de código aberto StarUML com sua extensão espacial foi utilizada para representar as características geográficas, para assim obter o modelo conceitual, adquirindo as características que possibilitaram a modelagem dos dados geográficos com as características da OMT-G, seguindo assim o postulado da INDE. A ferramenta StarUML serviu para a visualização e análise de maneira mais adequada do modelo de dados geográficos, porém o processo para transformar este modelo em um banco de dados geográfico foi improdutivo, visto a inexistência de funcionalidades da ferramenta de modelagem geográfica StarUML. Sendo assim foi necessária a utilização de uma segunda ferramenta especializada em modelagem de banco de dados.

Para adquirir eficiência e conseguir as abstrações necessárias durante o processo de modelagem do banco de dados geográfico, a ferramenta DbVisualizer foi utilizada para manter e atualizar o modelo nos níveis lógico e físico. Este software possui um conjunto de funcionalidades capaz de representar as características objeto-relacionais, ficando as representações geográficas no modelo produzido pelo StarUML e extensão OMT-G. Utilizando o DbVisualizer houve a possibilidade de, através do modelo de banco de dados, gerar o

banco de dados no SGBD PostgreSQL efetivamente, obtendo scripts SQL ou realizando uma conexão diretamente no Sistema Gerenciador de Banco de Dados (SGBD).

Assim, a ferramenta StarUML foi utilizada numa primeira etapa para gerar o modelo conceitual, visto que permite representar características geográficas mediante sua extensão espacial. Dada a limitação do StarUML para gerar o modelo lógico, foi utilizado a ferramenta DbVisualizer para obter o modelo lógico na segunda etapa.

3.2 Sistema Gerenciador de Banco de Dados

O PostgreSQL é um sistema gerenciador de banco de dados objeto relacional, gratuito e de código fonte aberto na sua versão 9.2. O PostgreSQL possui total compatibilidade com conceitos de ACID (acrônimo de Autenticidade, Consistência, Isolamento e Durabilidade), este é o conceito definido na computação para a caracterização da transação, suporte a chaves estrangeiras, junções, visões, gatilhos e procedimentos armazenados, possui compatibilidade com os principais tipos de dados definidos na ISSO SQL:1999.

Usado em projetos corporativos, PostgreSQL é dotado de funcionalidades sofisticadas que possibilitam o controle de concorrência multiversionado, recuperação dos dados em um ponto no tempo, cópias de segurança, um planejador de consultas e registrador de transações sequencial para tolerância a falhas.

3.3 Extensões Espaciais

O PostgreSQL descrito acima não tem a capacidade de armazenar e manipular dados geográficos de forma adequada, para isso é necessária uma extensão com tal capacidade. A diferença entre uma entidade geográfica e não geográfico num determinado banco de dados é apenas um atributo que tem em sua definição as características necessárias para armazenar um dado na forma de vetores, assim como a característica geográfica que posiciona aquele conjunto de vetores no plano terrestre, utilizando uma definição cartográfica específica, quer dizer o sistema de referência espacial.

Neste trabalho foi utilizada a extensão espacial do PostgreSQL de código aberto, denominado PostGIS.

3.4 Servidor de Serviços Geográficos

É uma componente tecnológica que constitui parte do segmento de uma arquitetura de SIG, sua atribuição é fornecer através de um canal de comunicação funcionalidades de visualização e manipulação de dados geográficos vetoriais e matriciais. Atualmente, estes serviços são fornecidos por um computador servidor a outros computadores clientes numa rede de computadores local (intranet) ou global (internet). Um servidor de serviços geográficos é um software que em seu núcleo implementa funcionalidades capazes de administrar serviços geográficos. Neste trabalho foi utilizado o servidor GeoServer na sua versão 2.2.4.

O GeoServer é um servidor de código aberto escrito em Java que permite aos usuários compartilhar e editar dados geoespaciais. Usando padrões abertos definidos pelo OGC,

o GeoServer permite uma grande flexibilidade na elaboração dos mapas e compartilhamento dos dados. O GeoServer pode exibir dados em qualquer aplicativo de mapeamento, como Google Maps, Google Earth, Esri ArcGIS, etc [23]. Seu ponto forte é a interoperabilidade, o ponto negativo encontrado no software foi a pobre interface com o usuário para configuração de simbologia e o ainda inicial módulo de disponibilização de serviços web em padrão REST [23]. O GeoServer é um servidor de WMS, WCS e *Web Feature Service-Transaction* (WFS-T) [23].

Na implementação da aplicação, foram utilizados apenas os *Web Services* WMS e WFS. A maioria das camadas foram disponibilizadas em WMS, no formato de imagem em PNG. Apenas a camada de coordenadas foi disponibilizada em WFS, para tornar possível a obtenção de dados acerca dos pontos exibidos no mapa. A Figura 3.1 apresenta a interface padrão do GeoServer. Nessa figura, pode-se observar as camadas criadas no GeoServer com algumas informações importantes sobre elas. Vale citar a coluna “Todos los Formatos” (Todos os formatos), que exibe todos os formatos possíveis em que uma camada pode ser gerada. Como exemplo, tem-se para o WMS, os formatos GIF, GeoRSS, JPEG, PNG, OpenLayers, etc. Para o WFS, pode-se citar os formatos CSV, GML2, GeoJSON, Shapefile, etc.



Figura 3.1: Tela com as camadas existentes no GeoServer.

Uma funcionalidade interessante do GeoServer é a capacidade que ele possui de, recebendo dados de uma camada em um sistema geodésico de referência, como o SIRGAS 2000 (EPSG: 4674) por exemplo, projetá-la em uma outra projeção, como o SAD69 (EPSG:4618) por exemplo. Essa funcionalidade foi de grande utilidade para o projeto, haja vista que algumas camadas foram geradas na projeção SIRGAS 2000 e projetadas na aplicação com o sistema de coordenadas do GOOGLE (EPSG: 900913), que foi a projeção padrão utilizada pela plataforma do SIG implementado.

3.5 Quantum GIS

O QuantumGIS, ou QGIS, é um Sistema de Informação Geográfica de código aberto que executa tanto em Linux, Windows, MAC OSX, ou Android [24]. Essa ferramenta pode trabalhar tanto com dados vetoriais quanto com dados raster (imagens de satélite). Além de ter várias outras funcionalidades para análise espacial dos dados tais como cálculo da distância entre dois pontos no mapa. Para esse projeto, o QGIS foi utilizado basicamente

como uma ferramenta de teste da visualização dos dados espaciais que seriam colocados no sistema. Para realizar a visualização de dados mostrada na 3.2, primeiramente foi carregado o *shapefile* das regiões no QGIS. Esse mapa representa a primeira camada. Depois foram carregados os dados espaciais do banco de dados, isto é, os pontos das amostras coletadas de solos e vegetação para o QGIS, formando assim a segunda camada. Por último, essas camadas foram sobrepostas para formar a Figura 3.2.

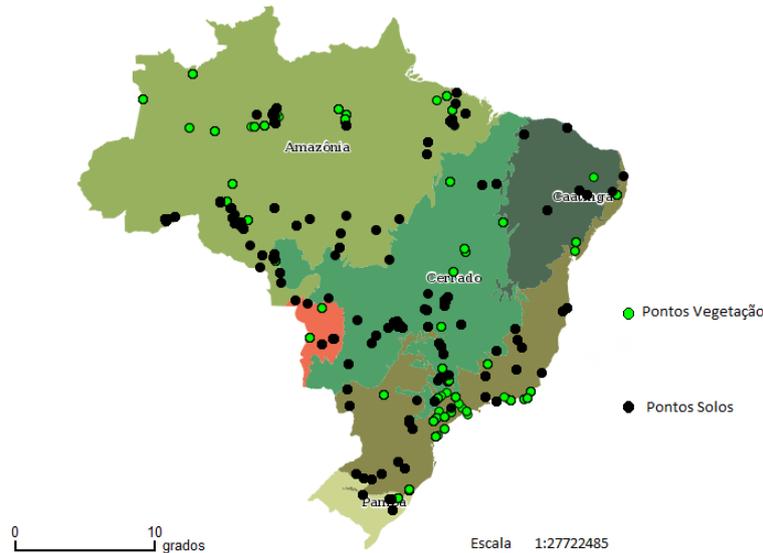


Figura 3.2: Visualização dos pontos de solos e vegetação no mapa dos Biomas de Brasil. Fonte própria.

O QGIS também permite uma organização dos dados de modo a facilitar a interpretação das informações a partir de uma simples visualização.

3.6 GeoKettle

Para carregar os dados no banco de dados foi utilizada uma ferramenta ETL (*Extract Transform Load*) o Geokettle mostrada na Figura 3.3.

GeoKettle na sua versão 2.5 é uma poderosa ferramenta ETL, orientada para metadados espaciais dedicado à integração de diferentes fontes de dados espaciais para a construção de estoques e atualização de dados geoespaciais. GeoKettle permite a extração de dados, a transformação de dados, a fim de corrigir erros, fazer um pouco de limpeza de dados, alterar a estrutura de dados, torná-los compatíveis com os padrões definidos, e o carregamento de dados transformados em um SGDB em modo OLTP ou OLAP / SOLAP, arquivo SIG ou *Web Services Geoespacial* [25].

Uma das dificuldades encontradas para a extração de dados foi a impossibilidade de se fazer uma extração da planilha como um todo por questões de restrição de integridade no banco gerado a partir do modelo. Assim para cada tabela do modelo relacional, foi necessária a realização de uma transformação diferente.



Figura 3.3: Ferramenta de extração, carregamento e transformação de dados Geokettle [25].

3.7 OpenScales

OpenScales é um framework de mapas de código aberto (LGPL) escrito em *ActionScript 3* e *Flex*, permite construir aplicações ricas para internet (RIA) de mapas.

OpenScales foi projetado para executar em todos os tipos de dispositivos:

- *Mobile (ActionScript 3)*: Testado em *Android Flash Player 10.1*.
- *Web (Flex)*: *OpenScales* roda em *Adobe Flash Player 10* ou superior, cobrindo 98% dos atuais *web browsers* sob Windows, Mac e Linux.
- *Desktop (AIR)*: Aplicação de mapa para desktop baseada em *OpenScales* e AIR pode ser útil para gerenciamento de modo off-line, abrir arquivos locais como shapefiles.

Algumas das principais características do *framework OpenScales* são: suporte à diferentes tipos de camadas, protocolos OGC: WMS/WMS-C 1.1.1, WFS 1.0, suporte a *GetCapabilities*, etc.

3.8 XGIS

XGIS é um software livre, construído sobre o *framework* de mapas *OpenScales*. Ele é um Sistema de Informação Geográfica implementado em *Flex* e *ActionScript 3* que oferece uma interface bastante amigável, com várias funcionalidades, por exemplo, a possibilidade de se carregar um *shapefile* dinamicamente, a possibilidade de exportar camadas existentes na aplicação, ferramentas de medição de distâncias sobre uma camada, etc.

O XGIS é uma contribuição do aluno do Instituto de Geociência Mariano Pascual [26] [27]. A aplicação é bastante customizável, o que corrobora a sua escolha para utilização nesse projeto. Algumas das vantagens em relação ao i3Geo por exemplo, é o fato de o XGIS possuir uma interface bastante amigável, de fácil navegação e utilização, e o mais importante é que a aplicação não necessita realizar uma busca à base de dados cada vez

que o mapa é movimentado, pois as imagens das camadas são armazenadas em cache e a aplicação carrega uma área da imagem maior do que a área que está sendo visualizada, o que diminui o número de carregamentos das imagens.

3.9 Arquitetura Proposta para Integração de Dados Espaciais

Este trabalho tem como proposta definir uma arquitetura de integração de dados espaciais distribuídos em múltiplos provedores de serviços relativos a amostras de solos e vegetação coletados nos biomas brasileiros, com a finalidade de compartilhar e reutilizar dados espaciais, tornar o sistema mais flexível às mudanças e dar respostas ágeis. A Figura 3.4 apresenta a arquitetura proposta baseada na arquitetura SOA-SDI apresentada na seção 2.5 do capítulo 2, nas recomendações da INDE, nos padrões OGC e utilizando software livre. A arquitetura proposta é composta de três camadas: a camada de provedores de informação geográfica, a camada de integração e a camada de apresentação.

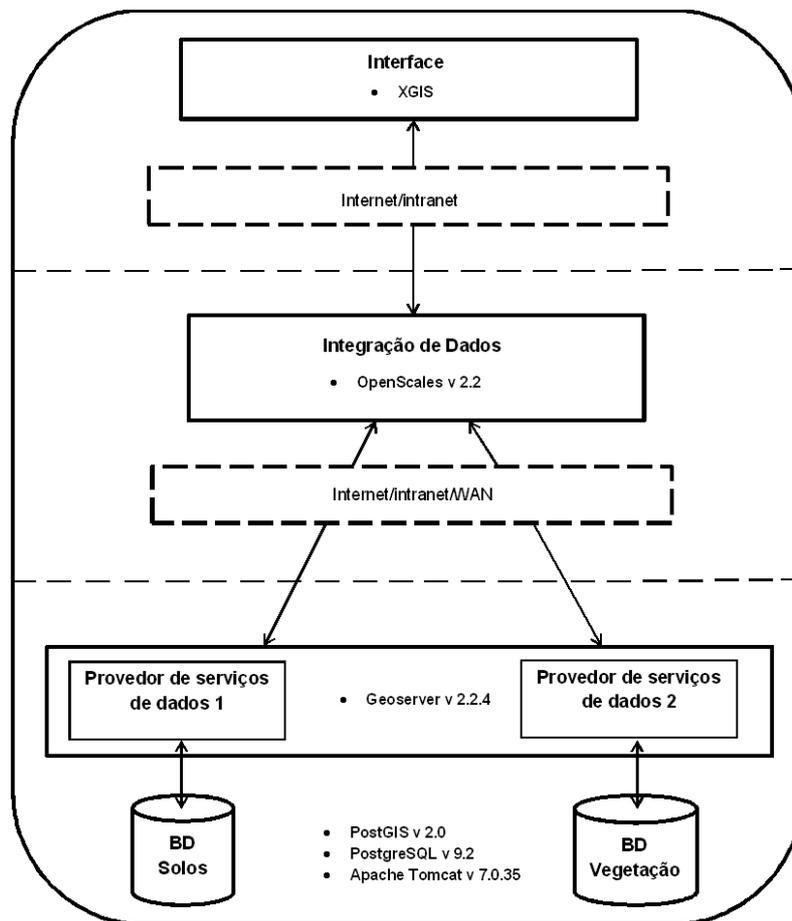


Figura 3.4: Arquitetura Proposta para Integração de Dados Geoespaciais.

A camada de provedores de informação geográfica esta composta pelos bancos de dados espaciais de solos e vegetação. Como repositório de dados foi escolhido o PostgreSQL com

seu modulo geográfico PostGIS. Para implementação dos serviços WMS e WFS, a opção escolhida foi o Geoserver.

A camada de integração de dados combina todos os serviços de provedores de dados disponíveis, passa solicitação de dados de aplicativos do cliente para os serviços de provedor de dados e inversamente envia de volta os dados de resposta dos serviços de dados para aplicativos dos clientes. A ferramenta utilizada para esta tarefa foi OpenScales.

Finalmente a camada de apresentação tem a função de apresentar o resultado dos dados geográficos permitindo aos usuários finais interagir com a camada inferior a ferramenta escolhida foi XGIS que é um framework baseado em OpenScales permitindo aos desenvolvedores construir aplicativos ricos de mapeamento em internet.

Capítulo 4

Resultados

Este capítulo apresenta os resultados da pesquisa, sendo organizado da seguinte forma: a Seção 4.1 apresenta a consolidação dos modelos de dados geográficos. Na Seção 4.2 são apresentados as relações entre as distintas entidades identificadas no modelo. A Seção 4.3 apresenta a construção do banco de dados geográficos. A Seção 4.4 apresenta o processo de carregamento dos dados nos bancos de dados. Na Seção 4.5 são vistos os modelos de bancos de dados finais. Já na Seção 4.6 é apresentado o SIG para demonstração do trabalho.

4.1 Modelo de Banco de Dados Geográfico

4.1.1 Temas Relevantes Abordados no Processo de Modelagem

Ao alterar a temperatura e a distribuição de chuvas no planeta, as iminentes mudanças climáticas modificarão a distribuição dos seres vivos e o funcionamento de ecossistemas, e conseqüentemente, os ciclos biogeoquímicos. No Brasil, o desafio está em prever como as mudanças climáticas afetarão os principais biomas, dentre eles o Cerrado e Floresta Amazônica.

O presente estudo de caso está vinculado ao Programa de Pós-Graduação em Ecologia da Universidade de Brasília e à rede de pesquisa do “Instituto Nacional de Ciência e Tecnologia em Mudanças Climáticas” (INCT-CLIMA).

Os dados modelados são amostras coletadas ao longo dos biomas brasileiros georreferenciadas mediante longitude e latitude de folhas de plantas e solos.

4.1.2 Amostras de Vegetação e Solos

Uma amostra de solo consiste em uma pequena porção de terra capaz de representá-lo por meio de análises químicas e físicas. Já no caso da vegetação, a amostra consiste de coleta de folhas e haste de plantas. A diferença principal da coleta de amostras entre o solo e vegetação é que a coleta de amostras de solo é determinada por um horizonte que é um intervalo de profundidade no solo explicado adiante. Os processos de armazenamento, visualização e análise dos dados coletados pelos pesquisadores do Departamento de Ecologia da UnB, relativos à vegetação e solos, são atualmente complexos, por serem realizados com planilhas Excel, o que dificulta o manuseio e análise dos dados contidos

nelas. O preenchimento das planilhas foi feito manualmente, fato que fez com que alguns dados ficassem inconsistentes, dificultando também a extração e análise dos dados.

A Figura 4.1 mostra um fragmento da planilha de solos, onde se pode observar o grupo de dados relativo à localização geográfica da amostra que foi coletada. As duas colunas de latitude e longitude decimal contém duas casas decimais só fornecidas nas planilhas Excel.

C	D	E	F	G	H	I
Coordinates						
municipality	State	lat (S)	long (W)	dec_lat	dec_long	coord_got_at
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP
Brasília	DF	15° 57:00	47° 53:00	-15.95	-47.88	SP

Figura 4.1: Dados de Localização Geográfica da Planilha Solos.

Dados da Planilha de Solos

A planilha do solo contém mais de 2200 registros de amostras de solos e mais de 100 colunas de dados classificados e agrupados de acordo com a afinidade entre eles. Assim, a planilha está estruturada da seguinte forma:

- **Localização Geográfica:** Grupo de dados que descreve a localização exata onde as amostras foram coletadas. Esse grupo contém as colunas de Estado, Município, local, área local, local onde as coordenadas foram obtidas, coordenadas em graus e decimais. Para a análise do modelo somente foram consideradas coordenadas decimais já que as coordenadas em graus podem ser obtidas a partir das coordenadas decimais. A diferença entre as colunas *local* e *local onde as coordenadas foram obtidas* é basicamente que o *local* é o nome do local onde as amostras foram coletadas e *local onde as coordenadas foram obtidas* refere-se a um ponto central do município em inglês *Central Point of Municipality* (CPM) ou em algum lugar do local em inglês *Site Point* (SP).
- **Mudanças no uso da terra:** Esse grupo de dados descreve o uso e mudanças da terra através do tempo, contendo dados como a técnica de remoção, ano desde que a vegetação nativa foi removida, o tipo atual de uso da terra, o número de anos atual em que a terra foi mudada de uso, principais espécies no uso da terra, a vegetação predominante, tratamento da terra, manejo da terra e taxa de lotação.
- **Referências:** O grupo de dados das referências descreve basicamente a fonte da obtenção dos dados, tendo informações sobre as referências onde os dados foram coletadas, citação dos dados, algumas observações, URL se for o caso, e o tipo de referência que ela é (teses, artigos, revistas ou dados obtidos no campo).
- **Caracterização do local:** os dados desse grupo descreve o local da coleta, descrevendo informações como a altitude sobre o nível do mar, o bioma, o local onde a altitude

foi obtida que refere-se a um ponto central do município em inglês *Central Point of Municipality* (CPM) ou em algum lugar do local em inglês *Site Point* (SP), a classificação do clima *Koppen*, a classificação do solo brasileiro, norte-americano e FAO (*Food and Agriculture Organization of the United Nations*), a temperatura anual e a precipitação anual.

- Elementos químicos da amostra: neste grupo são classificados os elementos químicos estudados nas coletas, dentre eles estão o Alumínio, Ferro, Silício, Carbono, Nitrogênio, Fósforo, elementos compostos como $pH_CaCl_2_KCl$ e pH_H_2O .
- Horizonte: Um solo possui camadas horizontais de morfologia diferente entre si. Essas camadas são chamadas de horizontes. Na planilha, os horizontes estão representados mediante os intervalos de profundidade da amostra coletada. Para esse controle, tem as colunas profundidade inicial, profundidade final e composição da profundidade. A *profundidade composta* foi uma coluna adicionada com a finalidade de identificar as profundidades compostas a partir das simples.
- Existem dados livres, não agrupados como o conteúdo de argila, período da amostra, área plotada, ano da amostra, e o valor da densidade aparente entre outros.

Dados da Planilha de Vegetação

A planilha da vegetação contém mais de 6100 registros e mais de 100 colunas de dados agrupados e classificados de forma semelhante à planilha de solos. A planilha de vegetação contém grupos de dados similares aos da planilha de solo como a localização geográfica, as mudanças no uso da terra, as referências e a caracterização do local. A seguir são descritos os demais grupos de dados:

- Serapilheira: é a camada formada pela deposição e acúmulo de matéria orgânica morta em diferentes estágios de decomposição que reveste superficialmente o solo ou o sedimento aquático. Na planilha são descritos mediante o estudo de teor de N e P, taxa de decomposição de K, a queda total e a deposição de N e P.
- Morfometria: grupo de dados que descrevem a área foliar, a massa da folha seca, SLA (*Specific Leaf Area*), LMA (*Leaf Mass per Area*) e a espessura do tecido.
- Caracterização da espécie: esse grupo de dados descreve a forma da vida, se a amostra que foi coletada é legume ou não, o potencial de fixação de N, o grupo funcional, tipo de fenologia e a fixação N₂ estimada.
- Informação da amostra: grupo que contém informações a respeito de período da amostra, quantidade de amostra que foi coletada, área mostrada, ano da amostra, estação quando foi coletada a amostra.
- Descrição taxonômica: contém informações de gênero, família, epíteto, cf (incluindo aff quando o epíteto ¹ não é confirmado), variedade de subespécies e o nome comum. Além disso, contém a coluna autor onde armazena o nome do autor que descobriu e deu nome a taxonomia da amostra.

¹O nome específico em botânica é sempre a combinação do nome de gênero e o epíteto específico, exemplo, a espécie comumente conhecida como o pinheiro é "*Pinus palustris*", o nome do gênero é *Pinus* e o epíteto específico é *palustris*.

4.1.3 Modelos e Especificações para Modelagem de Banco de Dados Geográficos

A utilização do modelo de dados de solo existente em [28] foi fundamental para este trabalho, já que simplificou o custo temporal e de retrabalho no modelo de dados de solos, já que o modelo de solos foi baseada nesse trabalho. Por outro lado, com respeito ao documento das especificações da INDE, embora as normas e especificações abranjam categorias como Hidrografia, Relevo, Sistema de Transportes, Relevo e outras (visto no Capítulo 2 Seção 2.3.3), ela ainda apresenta limitações em algumas áreas de estudo como, por exemplo, a Ecologia.

SIG Solos

O modelo de dados do SIG Solos [28] é orientado aos pontos de amostragem, local onde são analisados os horizontes. Os horizontes de determinado amostra são analisados quanto a suas propriedades, física, química e mineralogia. No entanto, características relevantes para análise de solos não são completadas neste modelo, características como a caracterização do local, mudanças no uso da terra, já que não são informações inerentes aos solos.

Documento das Especificações da INDE

Cientes da limitação do documento das especificações da INDE para o caso de solos e vegetação especificou-se as tabelas de detalhes para as classes que não tinham especificações no documento da INDE, como uma contribuição ao documento das especificações técnicas da INDE, com a ajuda dos pesquisadores do Departamento de Ecologia, já que eles tinham maior conhecimento sobre as funcionalidades de cada atributo na planilha. Assim a Figura 4.2 mostra a especificação de atributos para a localização geográfica. As especificações de todos os elementos do modelo estão no Anexo A.

Classe		Descrição			Código			Primitiva Geométrica	
Localizacao_geografica	Informações da localização geográficas da amostra.				1.07			★	
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia		
local	Alfanumérico	30	O nome do local onde a amostragem foi executada	A ser preenchido	---	Nulo	---		
area_local	Real	-	Área do local em Hectares	A ser preenchido	---	Nulo	---		
Coordenadas_obtidas_em	Alfanumérico	10	Local onde as coordenadas foram obtidas	CPM	Ponto central do município	Nulo	---		
Decimal_latitude	Alfanumérico	20	Latitude decimal da amostra	SP	Em algum lugar do local	Nulo	---		
Decimal_Longitude	Alfanumérico	20	Longitude decimal da amostra	A ser preenchido	---	Não Nulo	---		
GeometriaAproximada	Boleano	-	Indica que a geometria adquirida é aproximada, em relação à escala prevista para o produto cartográfico.	Sim	---	Não Nulo	---		
				Não	---	Nulo	---		

Figura 4.2: Detalhes de desenvolvimento da classe Localizacao geografica.

A coluna atributo descreve o nome dos atributos da classe, a coluna tipo descreve o tipo de dado que será assinalado ao atributo, a coluna tamanho descreve o tamanho do tipo de dado caso tenha, na coluna descrição é colocada uma pequena descrição do atributo, já a coluna domínio apresenta o domínio do atributo, sendo especificada somente se existir, caso contrário coloca-se “ a ser preenchido”. A coluna seguinte é uma descrição para cada domínio do atributo se existir, a coluna requisito especifica se o atributo será nulo ou não, adicionalmente a coluna fotografia é mantida da versão original do documento das especificações, pois ela contém um hiperlink para imagens. O restante de classes e detalhes das especificações do modelo de dados do solo e da vegetação estão detalhadas no anexo “A”.

4.2 Relacionamentos dos Modelos de Dados

A integração das bases de dados acontece por meio dos relacionamentos existentes entre as entidades. Este relacionamento deve obedecer à semântica do contexto modelado, e regras de negócio.

Para o processo de integração das bases de dados, foram utilizados dois tipos de critérios de integridade dos relacionamentos, o relacional e o geográfico, ambos possuem características e aplicações diferentes.

No banco de dado dos solos houve relações geográficas e relações não geográficas, que integram as classes de localização geográfica à entidade Horizonte. Na Figura 4.3 pode-se observar as relações geográficas entre as entidades *Estado*, *Município* e *Localização geográfica*. As entidades geográficas estão representadas pelas primitivas geográficas do lado superior esquerdo, nas quais pode-se observar a representação de dois polígonos (entidades *Estado* e *Município*) e representação de um ponto (entidade *localização geográfica*), onde são armazenadas as coordenadas. As relações geográficas estão identificadas com linhas pontilhadas, e integra as entidades *Estado-Município* e *Município-Localizacao geográfica*, já uma relação não geográfica liga a entidade *localização geográfica* com a entidade *Horizonte*.

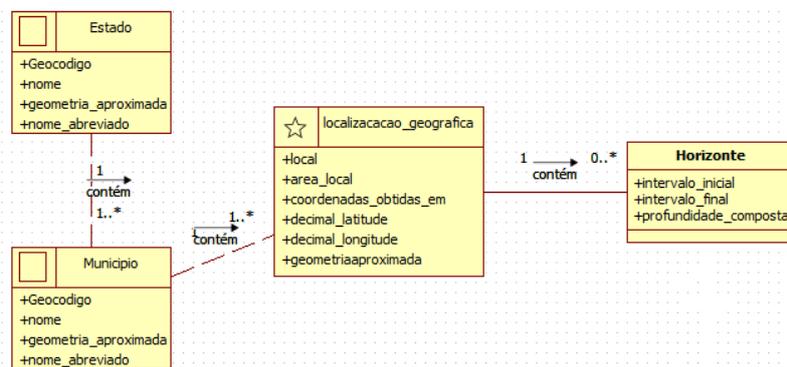


Figura 4.3: Relações Geográficas e não Geográficas no Modelo de dados do Solo.

Para o caso do banco de dado da vegetação, a Figura 4.4 mostra um fragmento do modelo de dados da vegetação. A estrutura é a mesma que no caso do modelo do solo

com a diferença que a entidade *localização geográfica* está integrada com a entidade central *Amostra Vegetação*.

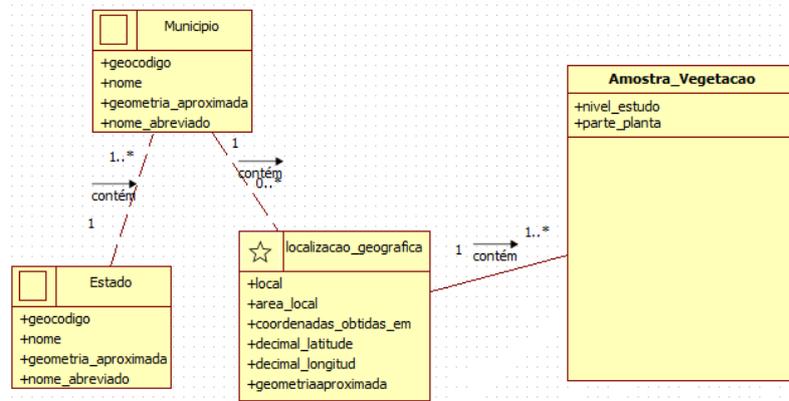


Figura 4.4: Relações Geográficas e não Geográficas, Modelo de dados da Vegetação.

Dos dois fragmentos pode-se observar a semelhança entre as entidades *Estado*, *Município* e *localizacao geográfica* nos modelos, por causa da adaptação dos modelos com as especificações técnicas da INDE para modelamento de dados. Assim atributos como *geocodigo* e *geometriaaproximada* foram adicionadas seguindo a padronização das especificações. Para as limitações do documento foi realizado um documento com as especificações para o resto das entidades, como se mostra a continuação.

Nas Figuras 4.5 e 4.6 pode-se observar a semelhança entre as entidades *caracterização local*, *Classificacao solo* e *Temperatura anual*. Em ambos os casos e integradas às entidades centrais de *amostra vegetacao* e *amostra solo*. Para isso foram utilizadas as especificações de atributos feitas, para padroniza-los e uniformiza-los, daí a semelhança nos nomes dos atributos. Por exemplo, para a entidade *Temperatura anual* do modelo de vegetação, a Figura 4.7 mostra as especificações para cada atributo, assim o atributo *temperatura minima* será do tipo real e não nulo.

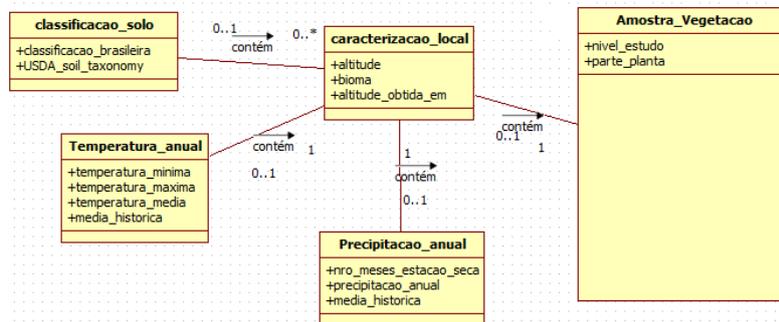


Figura 4.5: Fragmento Modelo de Dados vegetação.

No caso da entidade *precipitacao anual*, os atributos no modelo de solos e vegetação são distintos.

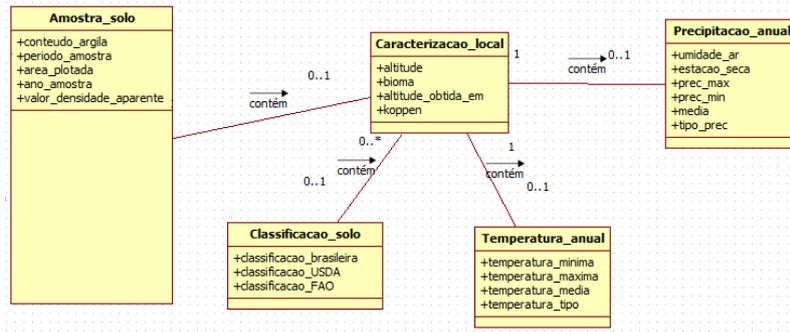


Figura 4.6: Fragmento Modelo de Dados de Solos.

Dessa forma cada uma das entidades nos modelos foi especificada com o fim de uniformizar e diminuir a heterogeneidade semântica dos atributos.

Classe	Descrição			Código	Primitiva Geométrica		
	Descreve as informações relacionadas à temperatura anual do local estudado.				Descrição	Requisito	Fotografia
Temperatura_anual				1.09			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Temperatura_minima	Real	-	Indica a temperatura mínima anual média do ar em graus célsius.	A ser preenchido	---	Nulo	
Temperatura_maxima	Real	-	Indica a temperatura máxima anual média do ar em graus célsius.	A ser preenchido	---	Nulo	---
Temperatura_media	Real	-	Indica a temperatura média anual do ar em graus célsius.	A ser preenchido	---	Nulo	---
Media_historica	Real	-	Temperatura do ar média anual para uma série histórico específica.	A ser preenchido	---	Nulo	---

Figura 4.7: Especificações dos Atributos Entidade Temperatura anual Modelo de dados da Vegetação.

A Figura 4.8 mostra a integração das entidades *Horizonte* e *Química amostra* mediante uma relação de um a muitos. A entidade *Horizonte* representa a profundidade em que a amostra foi coletada. Neste caso em centímetros representado pelos atributos intervalo inicial e final, o atributo *profundidade composta* descreve se o intervalo em questão é a soma de dois ou mais intervalos anteriores.

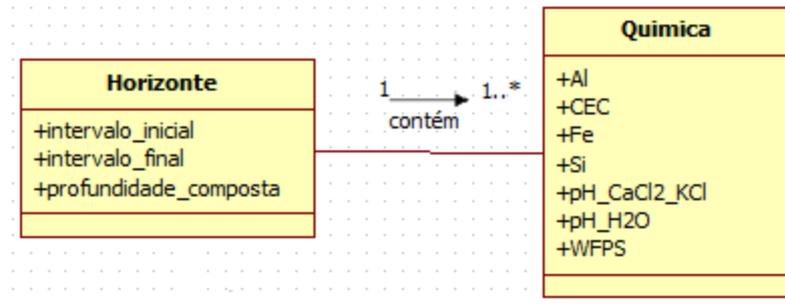


Figura 4.8: Integração Entidades Horizonte - Química amostra Modelo de Dados do Solo.

A entidade *Química amostra* contém os elementos químicos estudados e analisados para uma amostra dentro deles o Alumínio, Ferro, silício. Ainda a entidade *Química amostra* contém informações do Carbono, Nitrogênio e Fosforo, só que elas são integradas às entidades respectivas mostradas na Figura 4.9.

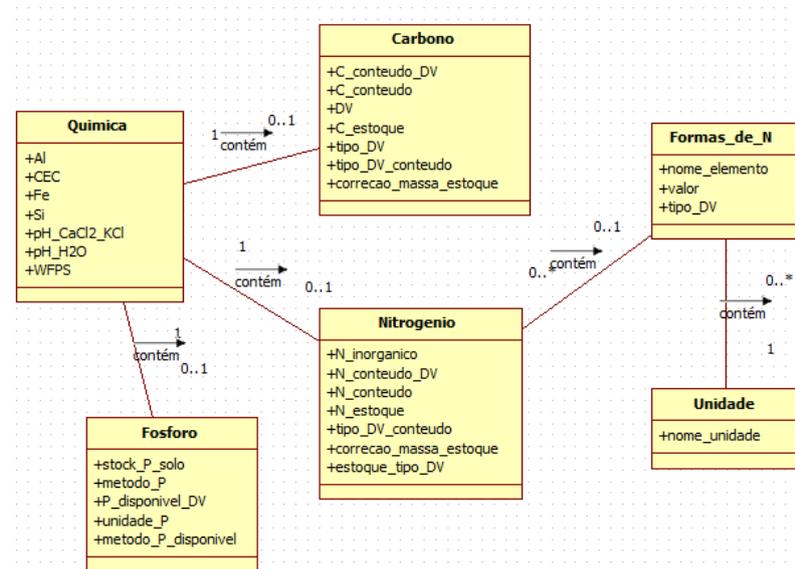


Figura 4.9: Entidades Fosforo, Nitrogenio, Carbono Integradas a Entidade Química Modelo de Dados Solos.

Foram criadas novas entidades para *Carbono*, *Nitrogênio* e *Fosforo* pois cada uma delas tem a própria estrutura. A entidade *Nitrogênio* está integrada também a uma nova entidade, *Formas de N*, que contém informações relacionadas ao $N - NO_3$, $N - NH_4$, $d15N - NO_3$, $d15N - NH_4$, $d15N$ e DIN . Cada uma delas com uma unidade de medida

diferente, razão pela qual criou-se a entidade *Unidade* que contém as diferentes unidades de medida utilizadas na planilha.

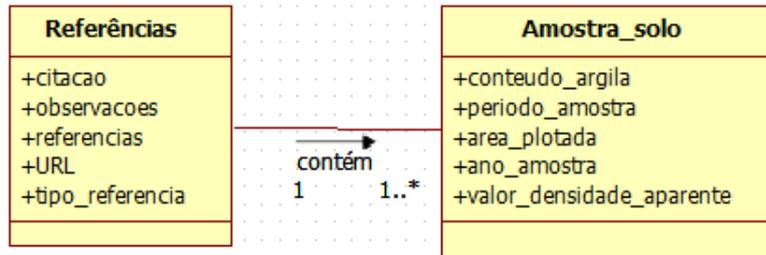


Figura 4.10: Integração Entidades *Referências* e *Amostra solos* Modelo de Dados Solos.

A Figura 4.10 mostra as entidades *Referências* e *Amostra solo* integradas mediante uma relação de um a muitos. Cada amostra coletada tem uma referência para saber a origem da informação (teses, artigos, revistas) ou se foi coletada no campo mesmo.

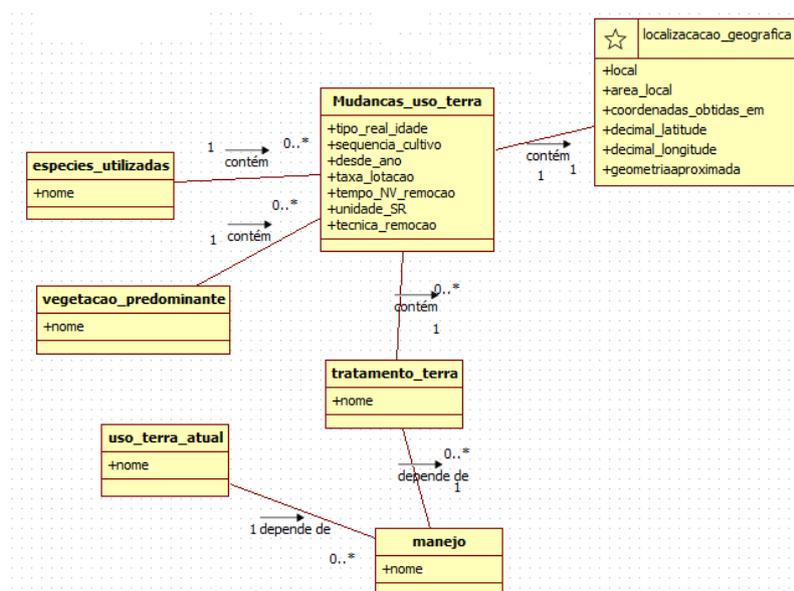


Figura 4.11: Entidade *Mudancas uso terra* integrada a *Localização geográfica*.

Na Figura 4.11 pode-se observar a integração da entidade *Mudancas uso terra* com a entidade *Localização Geográfica*. A entidade sobre as mudanças do uso da terra contém as informações do uso e mudanças na terra ao longo do tempo, então contém atributos que descrevem a técnica de remoção, taxa de lotação, sequência do cultivo entre outros. Contém também atributos que viraram novas entidades como espécies utilizadas, a vegetação predominante, tratamento da terra, o manejo e o uso da terra atual.

Existe uma relação de dependência entre as entidades *uso terra atual*, *manejo* e *tratamento terra*. Assim, uma terra usada como vegetação nativa (*uso terra atual*), pode ser queimada, fertilizada ou irrigada (*manejo*). Se for queimada, requerem-se a informação

se ela foi queimada acidentalmente, queimada com controle ou se simplesmente não foi queimada ainda (*tratamento terra*).

Por outro lado, algumas das relações mais importantes no modelo da vegetação são detalhadas a seguir.

A Figura 4.12 mostra a integração das entidades simples à entidade central:

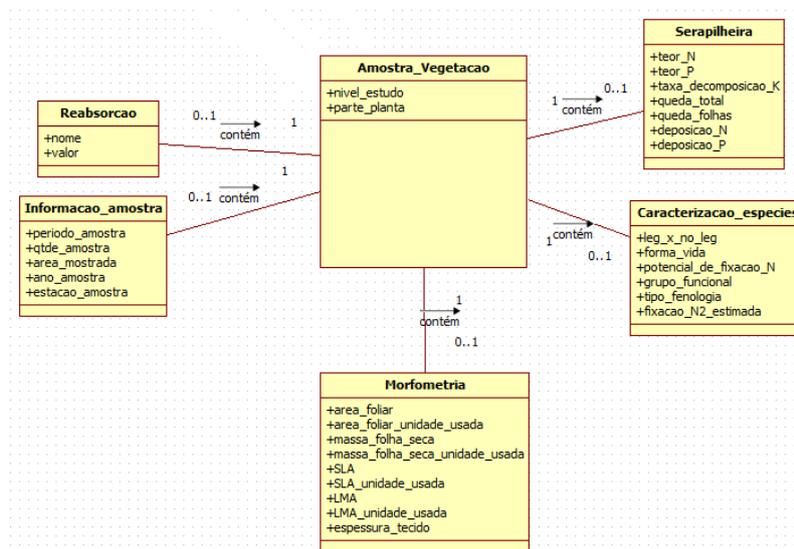


Figura 4.12: Integrações com Entidades Simples com a Entidade Central *Amostra vegetacao*.

- Reabsorção que descreve a percentagem de reabsorção, é baseada em variações na concentração de Nitrogênio e Fosforo por unidade de massa foliar.
- Serapilheira descreve a camada formada pela deposição e acúmulo de matéria orgânica morta em diferentes estágios de decomposição que reveste superficialmente o solo. Integrada com uma relação simples de zero a um.
- A entidade *informacao amostra* está integrada com a entidade central com uma relação de zero a um, contendo informações básicas como área mostrada, ano da amostra, estação em que foi pega, entre outras.
- Morfometria entidade onde são armazenadas as informações sobre área foliar, a massa da folha seca, entre outras. Integrada com uma relação simples de zero a um.
- A entidade *caracterizacao especies* contem informações sobre a forma da vida, potencial de fixação do Nitrogênio, o grupo funcional ao qual pertence, entre outros.

Por último, a entidade *descricao taxonomica* Figura 4.13 é integrada com uma relação simples de zero a um com a entidade central. Existem dependências integradas à entidade *descricao taxonomica* que são o gênero e a família da espécie estudada. Adicionalmente a entidade autor descreve o nome do autor que descreveu a espécie.

Assim os modelos conceituais para solos e vegetação foram criados e podem ser observados no anexo “B”.

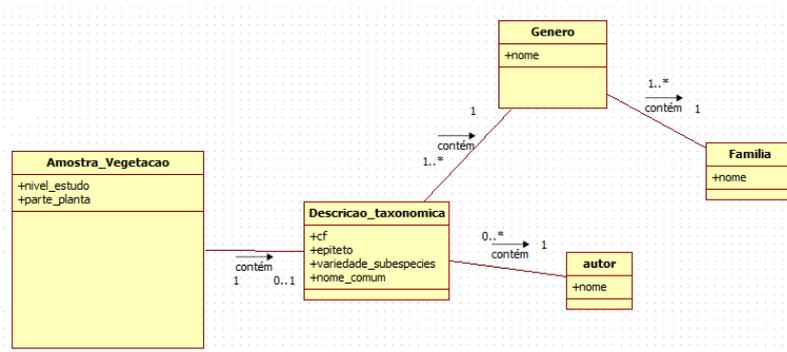


Figura 4.13: Integração entidade *Descricao Taxonomica* e *Amostra Vegetacao*.

4.3 Construção dos Bancos de Dados

Uma vez consolidado o modelo conceitual, esta seção tratará dos aspectos de criação dos bancos de dados.

4.3.1 Criação de Modelo Lógico

O modelo lógico é parte do processo de criação de um banco de dados, ele é baseado no modelo conceitual apresentado na seção 4.2, sendo o modelo conceitual utilizado como parâmetro de entrada nesta fase. Todos os atributos das diversas classes são identificados, os nomes padronizados e recebem uma tipificação, quer dizer, o tipo de dado que esta propriedade será capaz de armazenar.

As Figuras 4.14 e 4.15 mostram os modelos lógicos para Solos e Vegetação.

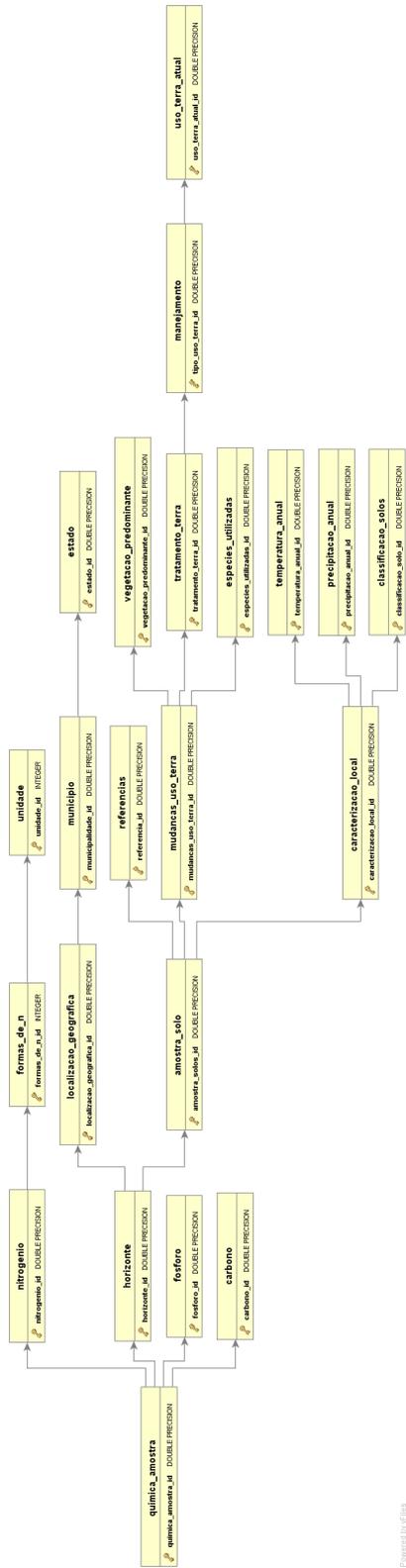
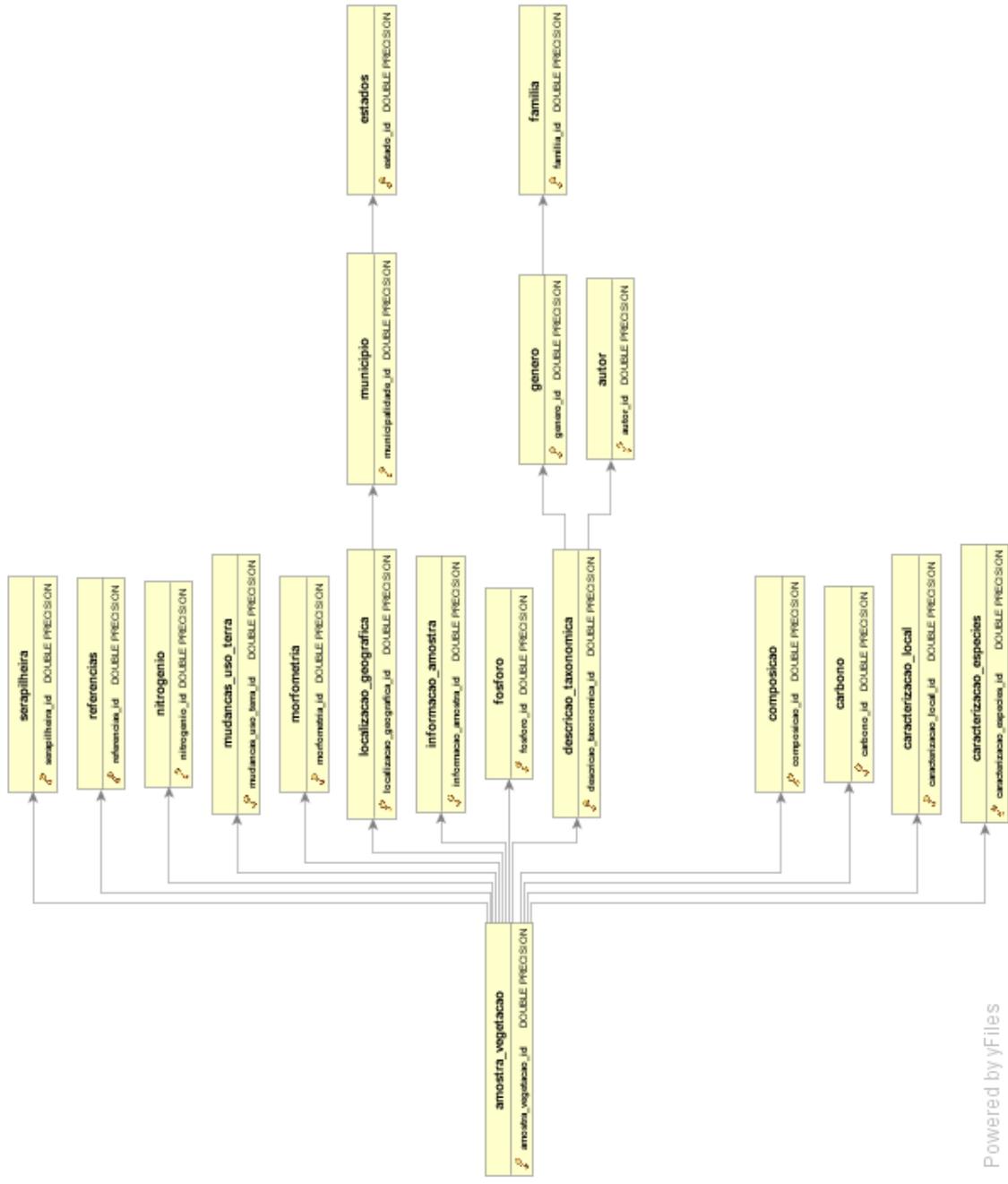


Figura 4.14: Modelo de Datos Lógico do Solo.



Powered by yFiles

Figura 4.15: Modelo de Dados Lógico da Vegetação.

4.3.2 Criação de Modelo Físico

O modelo físico aborda temas inerentes à tecnologia de banco de dados. Este modelo é gerado, no presente trabalho, para o SGBD PostgreSQL na sua versão 9.2. A finalidade desta fase é de transformar o modelo lógico ao modelo físico de acordo com a tecnologia do SGBD, por exemplo PostgreSQL. Esta fase deverá ser realizada toda vez que exista alguma alteração na tecnologia do SGBD.

4.3.3 Criação do *Script* de Construção do Banco de Dados

O *Script* de construção do banco de dados é gerado pelo processamento do modelo físico em linhas de comando denominadas *Structure Query Language* (SQL). Esta linguagem estruturada de consulta tem a capacidade de criar efetivamente o banco de dados com todas as suas entidades, atributos e relacionamentos. Feito isso, o SGBD neste caso o PostgreSQL agora é o responsável por armazenar e disponibilizar um conjunto de ferramentas para sua consulta e administração.

É importante observar que este script SQL, assim como o modelo físico, só funcionará efetivamente em determinada tecnologia de banco de dados, para a utilização em outra tecnologia é necessária a mudança do modelo físico, podendo ser feita utilizando-se de uma ferramenta de apoio a modelagem, por exemplo StarUML.

Um trabalho importante no *Script* é a geração do banco de dados, com as componentes geográficas, pois a ferramenta utilizada ainda não é integrada ao processo de concepção do banco de dados geográficos. Ao utilizar a extensão espacial PostGIS na sua versão 2.0, foi necessária adicionar uma linha de comando para cada entidade geográfica, com o objetivo de adicionar o atributo geográfico de acordo com sua especificação do modelo físico.

Sintaxe: *SelectAddGeometryColumn*(<nome da tabela>, <nome da coluna geográfica>, <identificador da referência espacial>, <tipo da geometria (ponto, linha, etc.)>, <dimensão da geometria (2D, 3D, etc.)>);

Para o caso da localização geográfica, tanto no banco de dados de solos e vegetação o *Script* foi:

```
selectAddGeometryColumn('localizacao geografica','theGeom',4674,'POINT',2);
```

Consolidado o *script* com as definições de comando para criação do banco de dados propriamente dito, tabelas, atributos descritivos e geográficos e relacionamentos, é necessária a utilização da ferramenta do SGBD denominada de PAdmin III (Postgres Administrator). Esta ferramenta possui dentre as funcionalidades, uma que processa *Scripts* SQL, que avalia sintaticamente toda a estrutura de comandos, onde, se houver algum problema identificado, o processo é interrompido até que a sintaxe seja corrigida. Quando a sintaxe estiver sem problemas, os comandos são executados sequencialmente, de forma a realizar cada uma das suas funções específicas.

4.4 Carga de Dados nos Bancos de Dados Geográficos

Com a estrutura de banco de dados criada no SGBD, o passo seguinte realiza a carga dos dados das planilhas para os bancos de dados. Esta seção descrevera o processo de carga dos dados na sua forma bruta, quer dizer, como encontrado nas planilhas fornecidas,

aos bancos de dados. Para isso se descreve o procedimento de carga com a ferramenta Geokettle, mostrado na Figura 4.16.

Uma das dificuldades encontrada para a extração de dados foi a impossibilidade de fazer uma extração da planilha inteira por questões de restrição de integridade no banco, gerado a partir do modelo. Assim para cada tabela do modelo relacional, foi necessária a realização de uma transformação diferente.

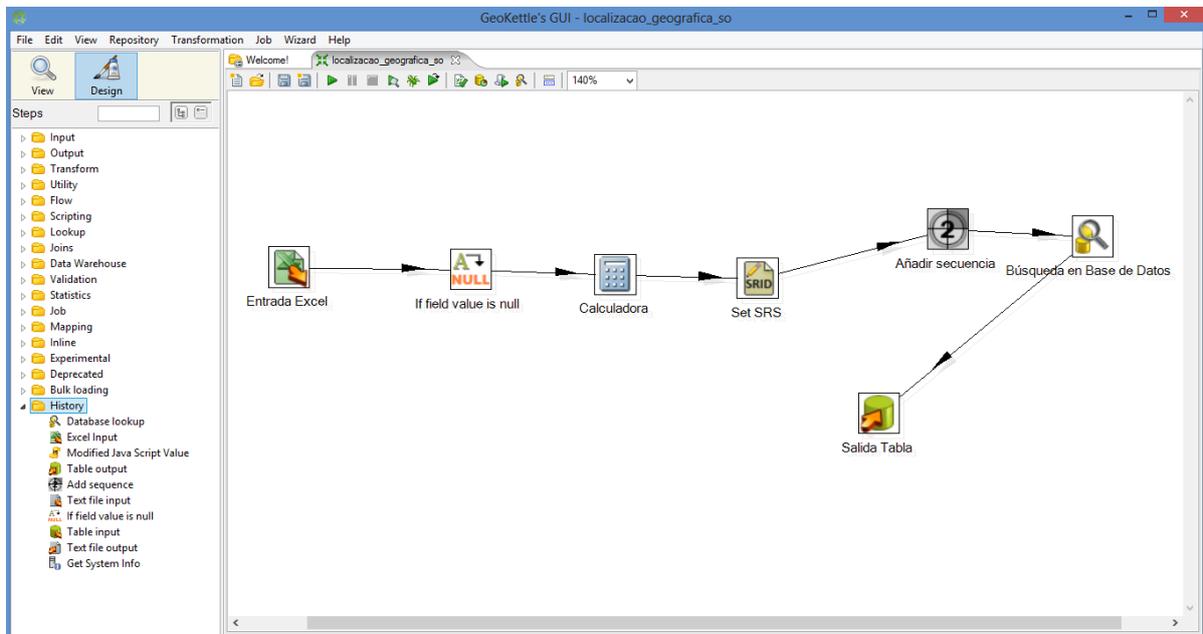


Figura 4.16: Processo de Extração de Dados da Entidade *Localizacao geografica* e Carregamento no Banco de Dados dos solos.

O primeiro passo é chamado de *Excel Input* onde as colunas desejadas para essa transformação são passadas como entrada e são tratadas com a função *trim* para eliminar os espaços no começo e final de cada registro. O próximo passo é a transformação dos registros nulos em *undetermined*, para que o banco não fique com registros vazios. Depois ocorre a transformação das coordenadas longitudinais em uma geometria do tipo ponto, graças ao passo *Calculador*. No próximo passo, *Set SRS*, é escolhido a projeção geográfica para a transformação das coordenadas longitudinais em geometria. Para esse projeto foi utilizada a projeção SIRGAS 2000 (EPSG 4674), pois este sistema foi oficialmente adotado pelo Brasil em 25 de Fevereiro de 2005 tornando-se o novo sistema de referência geodésico para o Sistema Geodésico Brasileiro (SGB) e para o Sistema Cartográfico Nacional (SCN).

O passo seguinte do processo ETL foi adicionar uma sequência auto incrementável, que foi a ID da tabela, utilizada como a chave primária da tabela *localizacao geografica*.

O próximo passo, *Database Lookup*, é referente ao processo de achar os respectivos índices para as chaves estrangeiras de cada registro. Nesse passo os dados do fluxo são comparados com os dados da tabela com a qual *localizacao geografica* tem um relacionamento. Quando ocorre um casamento de valores, a chave primária da tabela estrangeira é colocada na chave estrangeira da tabela *localizacao geografica*. Contudo há casos em

que a tabela estrangeira contém registros vazios. E o último passo, *Table Output*, carrega todos os dados do fluxo no banco. Para isso é preciso antes criar a tabela através de uma query SQL DDL (*Data Definition Language*) gerada pelo próprio passo, mas que pode ser editada pelo usuário. Com isso, é possível especificar quais serão as chaves primárias, as chaves estrangeiras e delimitar o tipo da variável em que o dado estará contido assim como o seu comprimento.

Ao final das etapas de extração, carregamento e transformação e carregamento dos dados, resta apenas verificar a integridade dos dados no banco PostgreSQL e em sua extensão PostGIS.

4.5 Modelo de Dados Final

Os modelos de dados resultantes deste trabalho foram a composição das relações apresentadas nas seções anteriores para atender às necessidades de análise e previsão. Os modelos produtos têm a capacidade de representar componentes geográficas vetoriais e atributos alfanuméricos, podendo esses variar no tempo. Desta forma é possível acompanhar as mudanças nas características dos objetos durante um período de tempo. Os modelos conceituais e lógicos dos bancos de dados encontram-se anexos a este trabalho.

Neste trabalho o banco de dados no SGBD PostgreSQL foi configurado para ser capaz de suportar tecnologia de extensão espacial, o PostGIS (OSGeo). Esta capacidade de multi-conexões permite o uso sincronizado de ferramentas distintas de consulta, edição e análise de dados geográficos que utilizam padrões abertos como os definidos pela OGC e outros proprietários.

4.6 Demonstração do Sistema de Informação Geográfica

A Figura 4.17 mostra a tela inicial do SIG, nela pode-se olhar os pontos de solos e vegetação ano Brasil inteiro. A configuração do servidor onde foi executada é a seguinte:

- Processador intel core i5
- Memória RAM 4GB
- Sistema Operacional Linux distribuição Ubuntu e Windows 8
- SGBD PostgreSQL v 9.2
- Extensão espacial PostGIS v 2.0
- Servidor de mapas Geoserver v 2.2.4

Como funcionalidades implementadas, podemos citar a exibição de dados obtidos a partir de clique com o mouse nos pontos da camada de Coordenadas. A Figura 4.18 mostra essa funcionalidade, que pode ser acessada após clique no botão do menu Navigation, seguido de clique em *Layers/Information*.

Na janela superior, com o nome de Layers, é possível gerenciar as camadas existentes na aplicação, podendo ativá-las e desativá-las, alterar transparência, mover para cima e



Figura 4.17: Tela Inicial do SIG.

para baixo e deletá-las. Logo abaixo, tem-se a janela Information, que exibe as espécies para um ponto clicado.

Como funções do próprio sistema XGIS temos ainda no menu *Navigation*, os botões *Add Layers Dynamically*, que permite adicionar camadas dinamicamente, nos formatos GPS, GeoRSS, KML, WFS, WMS e WMTS, *Total Extension*, que centraliza a tela no mapa do Brasil e *Overview*, que mostra uma janela no canto inferior direito para navegação dinâmica sobre o mapa mundial, como podemos visualizar na Figura 4.19.

Para o menu *Tools*, temos o botão *Export*, que permite exportar as camadas existentes na aplicação, nos formatos *Shapefile Zipado*, *KML*, *JPEG*, *PNG*, *PDF* e *OpenLayers*, nas projeções SIRGAS 2000 (EPSG:4674), SAD69 (EPSG:4618), WGS84 (EPSG:4326) e GOOGLE (EPSG:900913), como pode-se visualizar na Figura 4.20.

Em seguida, temos o botão *Measure*, que possui três ferramentas que permitem o cálculo de distâncias sobre um mapa. A primeira opção é medir distância, a segunda é medir distância angular e a terceira é medir área, como podemos observar na Figura 4.21.

Por último, o botão *GeoServer Console* (Figura 4.20), que redireciona o usuário para o painel de controle do GeoServer, permitindo o controle e gerenciamento das camadas existentes na aplicação.



Figura 4.18: Janela que exhibe as espécies para cada ponto.



Figura 4.19: Funções *Add Layers Dynamically*, *Total Extension* e *Overview*.



Figura 4.20: Função *Export*, utilizada para exportar camadas.

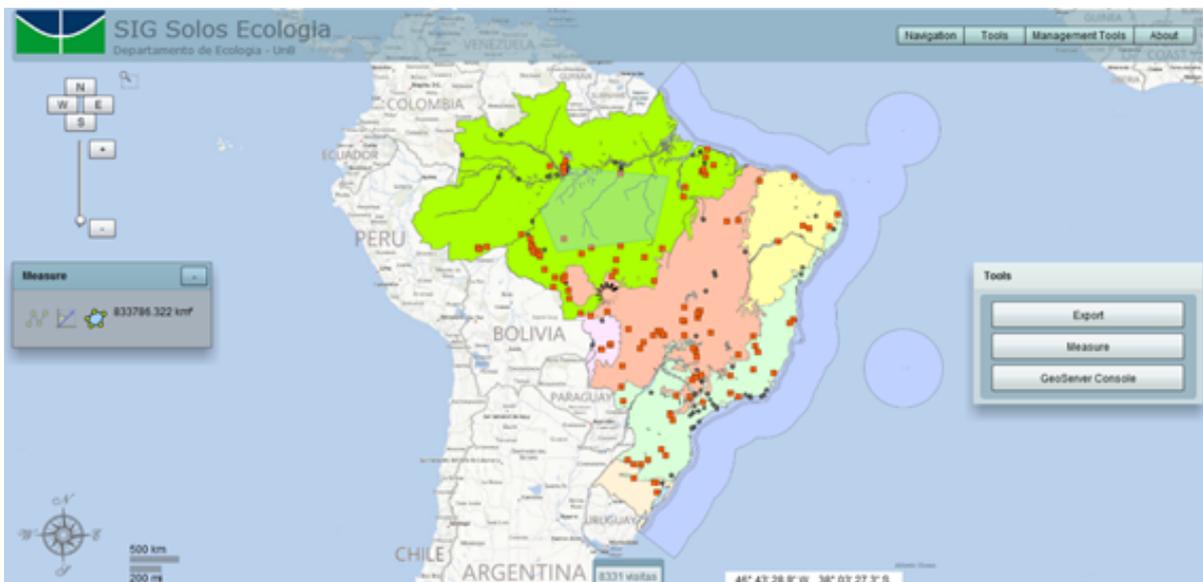


Figura 4.21: Ferramenta com funções de medições sobre o mapa.

Capítulo 5

Conclusões

Neste trabalho, foi proposta uma arquitetura orientada a serviços com *web services* para integrar e compartilhar dados espaciais do bioma brasileiro seguindo as recomendações da INDE.

Os processos de armazenamento, visualização e análise dos dados coletados pelos pesquisadores do Departamento de Ecologia da UnB, relativos à vegetação e solos, são atualmente complexos, por serem realizados com planilhas Excel, o que dificulta o manuseio e análise dos dados contidos nelas.

É fundamental para a construção de um sistema de informação que possa auxiliar o gestor da informação a obter conhecimento sobre o seu negócio, um banco de dados com informações consistentes. Na implementação de um banco de dados, o primeiro passo é o modelo de dados conceitual. Sendo assim, essa dissertação apresentou um modelo de dados de Solo e Vegetação dos Biomas Brasileiro.

O Brasil tem realizado esforços na criação de padrões e especificações para o desenvolvimento de aplicação com dados geográficos, como é o caso da INDE. Nesse contexto, o modelo conceitual dos dados foi baseado no OMT-G que é a modelagem definida pela INDE. A organização lógica dos dados de solos e vegetação permitirá também a produção e manutenção de novos dados, pois deverão seguir as diretrizes do modelo de especificação de banco de dados.

Um ponto interessante foi a utilização de ferramentas *open source*, como a ferramenta GeoKettle, o SGBD PostgreSQL com sua extensão espacial PostGIS, o servidor de mapas GeoServer, o *framework* de mapas OpenScales, que é um *framework* robusto para implementação de um SIG, o *framework* XGIS, desenvolvido pelo estudante de mestrado Mariano Pascual da UnB.

O *web Service* foi um componente importante no desenvolvimento de um SIG Web com SOA, uma vez que torna possível a integração de diferentes formatos de mapas interativos unificada em uma única imagem. Existem *web services* livres e pagos espalhados em instituições privadas e públicas, que o desenvolvedor pode utilizar ou compartilhar, com o fim de ter uma visão que permita ilustrar relações, conexões e padrões que não são facilmente visíveis em qualquer conjunto de dados, tornando mais fácil a tomada de decisões nas organizações. No projeto, os dados foram disponibilizados num servidor de Departamento de Geologia da UnB para o acesso externo dos dados.

Os produtos resultantes deste trabalho (banco de dados, SIG) agregam um valor final muito importante à manutenção e inspeção do equilíbrio ecológico das regiões estudadas

que cada vez mais sofre com ações antrópicas.

Pelos resultados obtidos o trabalho de análise de dados dos pesquisadores do Departamento de Ecologia da UnB foi aprimorado, os dados de solos e vegetação podem ser visualizados num mapa só, permitindo uma melhor análise e facilitando interpretação dos dados.

5.1 Contribuições

Segue as contribuições como fruto do trabalho da pesquisa:

- Publicação, e apresentação do artigo "Um SIG web com Integração de Base de Dados de Folha e Solo do Cerrado Brasileiro" na "7a conferencia Ibérica de Sistemas y Tecnologías de Información" realizada do 20 ao 23 de junho de 2012 em Madrid, España.
- Publicação do artigo "Uso de Serviços web em Sistemas de Informação Geográficos" no "IV Congreso Internacional de Computación y Telecomunicaciones" realizado em outubro de 2012, Lima, Perú.
- Desenvolvimento do sistema de informação Geográfica, "SIG Solos Ecologia".
- Extensão do documento das especificações Técnicas da INDE para o domínio de amostras de solos e vegetação dos biomas brasileiros.

5.2 Trabalhos Futuros

Como possíveis trabalhos futuros pode-se apontar:

- Criação de Ontologias relativas a amostras, para dar suporte à heterogeneidade semântica, ou seja, o problema da mesma informação ser armazenada de maneira diferente.
- Desenvolvimento de um CRUD (acrônimo de *Create, Read, Update e Delete* em língua Inglesa) para a administração dos dados no SIG.

Referências

- [1] Instituto Nacional de Pesquisas Espaciais, Brasil, *Tutorial sobre Bancos de Dados Geográficos*, *GeoBrasil*, 2006. x, 3, 4, 5, 6, 7, 8, 10, 13, 14, 15, 23, 24
- [2] G. M. Junior and H. Serrano, *Bancos de dados geográficos*. Escola Politécnica de Pernambuco.: Sumário. x, 10, 11, 12, 13
- [3] Comissão Nacional de Cartografia, Brasil, *Especificações técnicas para Estruturação de Dados Geoespaciais Digitais Vetoriais versão 2.0*, 2007. x, 16
- [4] Z. Sha and Y. Xie, “Design of service-oriented architecture for spatial data integration and its application in building web-based gis systems,” *Geo-Spatial Information Science*, vol. 13, Março 2010. x, 19, 20
- [5] F. Marcel and B. Martin, “Virtual database: Spatial analysis in a web-based data management system for distributed ecological data,” *Elsevier*, vol. 21, Novembro 2006. x, 21, 22
- [6] Comitê de Planejamento da Infraestrutura Nacional de Dados Espaciais, Brasil, *Plano de ação para implantação da Infraestrutura Nacional de Dados Espaciais INDE*, Janeiro 2010. x, 1, 8, 14, 15, 22, 23
- [7] Governo Brasileiro and Comite Executivo de Governo Eletronico, Brasil, *e-PING*, Dezembro 2010. 1, 19
- [8] Programa Nacional de Pós-Doutorado, Brasil, *Respostas às mudanças climáticas*., 2009. 2
- [9] I. Horrocks, “Ontology-Based Information Systems,” no. 1964, pp. 1–93, 2009. 3
- [10] T. Gruber, *Ontology*. I n the Encyclopedia of Database Systems: Springer-Verlag, 2009. 4
- [11] OGC, “Geography markup language.” <http://www.opengeospatial.org/standards/gml>. Acessado em Julho, 2012. 4, 17
- [12] S. Stevens, “On the theory of scales of measurement,” *Science Magazine*, no. 3, pp. 677–680, 1946. 5
- [13] R. Gutting, “An introduction to spatial database systems,” *The VLDB Journal—The International Journal on Very ldots*, vol. 3, no. 4, 1994. 9

- [14] M. Casanova, G. Câmara, C. Davis, L. Vinhas, and G. Ribeiro, *Banco de Dados Geográficos*. Curitiba: MundoGEO, 2005. 9
- [15] K. Borges, C. Davis, and A. Laender, “OMT-G: an object-oriented data model for geographic applications,” *GeoInformatica*, pp. 221–260, 2001. 10
- [16] M. Ortiz and G. Câmara, *Sistemas de Informação Geográfica para Aplicações Ambientais e Cadastrais: Uma Visão Geral*. 1998. 11, 24
- [17] Federal Geographic Data Committee, *A strategy for the national spatial infrastructure*, 1997. 14
- [18] The Australia New Zealand Land Information Council, Australia, *What is a asdi*, 2004. 14
- [19] pelo Instituto Geográfico Español and Universidade Politécnica de Madrid, España, *Curso sobre IDE*, 2008. 14
- [20] OGC, “Open geospatial consortium.” <http://www.opengeospatial.org/standards>. Acessado em Dezembro, 2012. 17
- [21] N. Bieberstein, S. Bose, M. Fiammante, K. Jones, and R. Shah, *Service-Oriented Architecture (SOA). Business Value, Planning, and enterprise Roadmap*. Escola Politécnica de Pernambuco.: FT Press, 2006. 18, 19
- [22] K. Elangovan, *GIS Fundamentals, applications and implementations*. New India publishing, 2006. 23
- [23] GeoServer, “Geoserver user manual.” <http://docs.geoserver.org/stable/en/user/index.html>. Acessado em Janeiro, 2013. 28
- [24] Quantum, “Quantum gis manual.” http://docs.qgis.org/user_guide/html/en/. Acessado em Janeiro, 2013. 28
- [25] Geokettle, “Geokettle documentation.” <http://wiki.spatialytics.org/doku.php?id=projects:geokettle>. Acessado em Janeiro, 2013. x, 29, 30
- [26] M. Pascual, “Desenvolvimento e implantação de sistema de informações geográficas disponibilizado pela internet para o instituto de geociências da universidade de brasília,” vol. 21, 2010. 30
- [27] M. Pascual, E. Alves, T. Almeida, G. França, H. Roig, and M. Holanda, “An architecture for geographic information systems on the web - webgis,” 2012. 30
- [28] S. Chagas, W. C. Junior, B. Bhering, A. Tanaka, and F. Baca, “Estrutura e organização do sistema de informações georreferenciadas de solos do brasil (sig solos – versão 1.0),” 2004. 36

Anexo “A”: Documento com as Especificações Técnicas para
Estruturação de
Dados Geoespaciais Digitais Vetoriais.

MODELO DE SOLOS

Classe	Descrição			Código	Primitiva Geométrica		
Referências	Referência completa onde a amostra foi pego.			1.01	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Citacao	Alfanumérico	30	Citação do documento. Ex. Lardy et al 2002	A ser preenchido	---	Nulo	---
Observacoes	Alfanumérico	150	Refere-se a algum comentário extra sob o trabalho	A ser preenchido	---	Nulo	---
Referencias	Alfanumérico	150	Referência completa do trabalho utilizado	A ser preenchido	---	Nulo	---
URL	Alfanumérico	300	URL onde foi encontrada a referência	A ser preenchido	---	Nulo	---
Tipo_Referencia	Alfanumérico	80	Tipo de referência	Artigo publicado	Artigos publicados em eventos científicos.	Nulo	---
				Dissertação	Dissertação de tese.		
				Capitulo de livro	Capitulo específico dum livro.		
				Dados não publicados	Dados que não foram publicados num evento científico.		
				No campo	Dados obtidos no campo mesmo		

Classe	Descrição			Código	Primitiva Geométrica		
Classificacao_Solo	Classificação do solo de acordo à nomenclatura brasileira, USDA e FAO (Food and Agriculture Organization of the United Nations)			1.02	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Classificacao_brasileira	Alfanumérico	100	Classificação do solo segundo a nomenclatura brasileira	A ser preenchido	---	Nulo	---
Classificacao_USDA	Alfanumérico	100	Classificação do solo segundo a	A ser	---	Nulo	---

			nomenclatura dos Estados Unidos	preenchido			
Classificacao_FAO	Alfanumérico	100	Classificação do solo segundo a nomenclatura da FAO	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Mudancas_Usso_Terra	Classificação do solo de acordo à nomenclatura brasileira, USDA e FAO (Food and Agriculture Organization of the United Nations)			1.03	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Tipo_real_idade	Inteiro	-	Numero de anos que a terra foi convertida.	A ser preenchido	---	Nulo	---
Sequencia_cultivo	Alfanumérico	150	Sequência de mudança no uso da terra	A ser preenchido	---	Nulo	---
Desde_ano	Dimensão Tempo	-	Ano em que iniciou o tipo de vegetação real	A ser preenchido	---	Nulo	---
Taxa_lotacao	Real	-	Densidade dos animais na pastagem	A ser preenchido	---	Nulo	---
Tempo_NV_remocao	Inteiro	-	Anos, desde que a vegetação nativa foi removida.	A ser preenchido	---	Nulo	---
Unidade_SR	Alfanumérico	10	Unidade de taxa de lotação	LU	Cabeça de gado por hectare	Nulo	---
				AH	Animais por hectare		
Tecnica_remocao	Alfanumérico	50	Técnicas de remoção de vegetação nativa	Queimada		Nulo	---
				Colheita			

				Não Lavrados			
				Cortada e queimada			

Classe	Descrição			Código	Primitiva Geométrica		
Manejo	Descreve o tipo de uso da terra			1.04	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumérico	100	Nome do tipo de uso da terra.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Uso_terra_atual	Descreve o tipo de uso atual da terra			1.05	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumérico	100	Nome do tipo de uso atual da terra.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Vegetacao_predominante	Descreve vegetação predominante na area estudada			1.06	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumérico	100	Nome da vegetação predominante na area estudada.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Especies_utilizadas	Principais espécies no uso da terra			1.07	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumérico	100	Nome das principais espécies no uso da terra.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
--------	-----------	--	--	--------	----------------------	--	--

Município, Estado	Pontos espaciais referenciais onde a amostra foi tomada			1.08	<input type="text"/>		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Geocodigo	Alfanumérico	7	Código criado pelo IBGE para identificar as unidades Político Administrativas da Divisão Territorial Brasileira e suas subdivisões operacionais (setor censitário), compondo chave única das unidades de coleta, apuração e disseminação de dados estatísticos e territoriais.	A ser preenchido	---	Nulo	---
Nome	Alfanumérico	80	Nome da municipalidade	A ser preenchido	---	Nulo	---
Geometria_Aproximada	Boleano	-	Indica que a geometria adquirida é aproximada, em relação à escala prevista para o produto cartográfico.	Sim	---	Não Nulo	---
				Não	---		
Nome_abreviado	Alfanumérico	20	Nome da municipalidade abreviado	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Localizacao_geografica	Informações da localização geográficas da amostra.			1.09			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
local	Alfanumérico	30	O nome do local onde a amostragem foi executada	A ser preenchido	---	Nulo	---
Área_local	Real	-	Área do local em Hectares	A ser preenchido	---	Nulo	---
Coordenadas_obtidas_em	Alfanumérico	10	Local onde as coordenadas foram obtidas	CPM	Ponto central do município	Nulo	---

				SP	Em algum lugar do local		
Decimal_latitude	Alfanumérico	20	Latitude decimal da amostra	A ser preenchido	---	Não Nulo	---
Decimal_Longitude	Alfanumérico	20	Longitude decimal da amostra	A ser preenchido	---	Não Nulo	---
GeometriaAproximada	Boleano	-	Indica que a geometria adquirida é aproximada, em relação à escala prevista para o produto cartográfico.	Sim	---	Não Nulo	---
				Não	---		

Classe	Descrição			Código	Primitiva Geométrica		
Temperatura_anual	Descreve as informações relacionadas à temperatura anual do local estudado.			1.10			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Temperatura_minima	Real	-	Indica a temperatura mínima anual média do ar em graus célsius.	A ser preenchido	---	Nulo	
Temperatura_maxima	Real	-	Indica a temperatura máxima anual média do ar em graus célsius.	A ser preenchido	---	Nulo	---
Temperatura_media	Real	-	Indica a temperatura média anual do ar em graus célsius.	A ser preenchido	---	Nulo	---
Temperatura_tipo	Alfanumérico	5	Indica o tipo de temperatura	CD	Dados coletados durante o estudo	Nulo	---
				HD	Dados históricos		

Classe	Descrição			Código	Primitiva Geométrica		
Precipitacao_anual	Descreve as informações relacionadas à precipitação anual do local estudado.			1.11			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Umidade_ar	Real	-	A média da humidade do ar anual para a região de amostragem em %.	A ser preenchido	---	Nulo	---
Estacao_seca	Inteiro	-	Número de meses por ano, que caracteriza a estação seca da região de amostragem.	A ser preenchido	---	Nulo	---
Prec_max	Real	-	Máximo de precipitação anual em milímetros.	A ser preenchido	---	Nulo	---
Prec_min	Real	-	Mínimo de precipitação anual em milímetros.	A ser preenchido	---	Nulo	---
Média	Real	-	Precipitação média anual, em milímetros.	A ser preenchido	---	Nulo	---
Tipo_prec	Alfanumérica	5	Indica o tipo de precipitação	CD	Dados coletados durante o estudo	Nulo	---
				HD	Dados históricos		

Classe	Descrição			Código	Primitiva Geométrica		
Caracterizacao_local	Classe que contém informações relativas à precipitação, temperatura, classificação do solo, altitude e bioma.			1.12			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Altitude	Real	-	Altitude respeito do nível do mar.	A ser preenchido	---	Nulo	---
Bioma	Alfanumérica	30	Nome do bioma brasileiro onde a amostra foi pego.	Amazônia	Constituída principalmente por uma floresta tropical	Nulo	---
				Cerrado	Constituído principalmente por savanas.		
				Mata Atlântica	Constituída principalmente por mata ao longo da costa litorânea que vai do Rio Grande do Norte ao Rio Grande do Sul.		
				Caatinga	Constituída principalmente por savana estépica.		
				Pampas	Constituído principalmente por vegetação campestre.		
				Pantanal	Constituído principalmente por savana estépica alagada em sua maior parte.		
Altitude_obtida_em	Alfanumérica	5	Local onde a altitude foi obtida.	CPM	Ponto central do município.	Nulo	---
				SP	Ponto do local		
Koppen	Alfanumérico	30	Descreve o sistema de classificação climática.	A ser preenchido.	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
fosforo	Descreve as informações relacionadas à temperatura anual do local estudado.			1.13			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Stock_P_solo	Real	-	Stock total de fosforo no intervalo de solo selecionado, em Kg ha-1.	A ser preenchido	---	Nulo	---
Metodo_P	Alfanumérico	30	Informação relacionada ao método de extração do P usada para determinar o fosforo disponível no solo.	A ser preenchido	---	Nulo	---
P_disponivel_DV	Real	-	Padrão de divergência média do fosforo disponível, em mg Kg-1.	A ser preenchido	---	Nulo	---
Unidade_P	Alfanumérico	30	Unidade de medida para o fosforo.	A ser preenchido	---	Nulo	---
Metodo_P_disponivel	Alfanumérico	30	Informação relacionada ao método de extração do P usada para determinar o fosforo disponível no solo.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Nitrogenio	Descreve o valor de N no solo em um intervalo de profundidade.			1.14			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
N_inorganico	Real	-	Soma total de N-NH4 e N-NO3 no solo amostrado.	A ser preenchido	---	Nulo	---
N_conteudo_DV	Real	-	O valor do erro padrão ou desvio para o N no intervalo do solo, em mg g-1.	A ser preenchido	---	Nulo	---
N_conteudo	Real	-	O conteúdo de Nitrogênio no intervalo de solo selecionado, em mg g-1.	A ser preenchido	---	Nulo	---
N_estoque	Real	-	Estoque de nitrogênio no intervalo de solo selecionado	A ser preenchido	---	Nulo	---
Tipo_DV_conteudo	Alfanumérico	5	Conteúdo tipo de dispersão estatístico.	SE	Erro standard	Nulo	---
				SD	Divergência standard		
Correcao_massa_estoque	Booleano	-	Correção de massa estoque.	True	Sim	Nulo	---
				False	Não		
Estoque_tipo_DV	Alfanumérico	5	Tipo de dispersão estatística do estoque.	SE	Erro standard	Nulo	---
				SD	Divergência standard		

Classe	Descrição			Código	Primitiva Geométrica		
Carbono	Descreve o valor de N total no solo em um intervalo de profundidade.			1.15			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
C_conteudo_DV	Real	-	O valor do erro padrão ou desvio para do C no intervalo do solo, em mg g-1.	A ser preenchido	---	Nulo	---

C_conteudo	Real	-	O conteúdo de Nitrogênio no intervalo de solo selecionado, em mg g-1.	A ser preenchido	---	Nulo	---
DV	Real	-	Erro padrão ou desvio para o C, em %.	A ser preenchido	---	Nulo	---
C_estoque	Real	-	Estoque de carbono no intervalo de solo selecionado, em Mg ha-1.	A ser preenchido	---	Nulo	---
Tipo_DV	Alfanumérico	5	Tipo de dispersão estatístico.	SE	Erro standard	Nulo	---
				SD	Divergência standard		
Tipo_DV_conteudo	Alfanumérico	5	Conteúdo tipo de dispersão estatístico.	SE	Erro standard	Nulo	---
				SD	Divergência standard		
				SD	Divergência standard		
Correcao_massa_estoque	Booleano	-	Correção de massa estoque.	True	Sim	Nulo	---
				False	Não		

Classe	Descrição			Código	Primitiva Geométrica		
Formas_de_N	Descreve as informações relacionadas ao N-NO3, N-NH4, d15N-NO3, d15N-NH4, d15N, DIN e d13C do local estudado.			1.16			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Valor	Real	-	O valor do erro padrão ou desvio para o solo N-NO3, N-NH4, d15N-NO3, d15N-NH4, d15N, DIN, d13C e densidade_aparente, em %.	A ser preenchido	---	Nulo	---

Tipo_DV	Alfanumérico	5	Tipo de dispersão estatístico.	SE	Erro standard	Nulo	---
				SD	Divergência standard		
Nome_elemento	Alfanumerico	20	Nome dos elementos químicos estudados	N-NO3	-	Nulo	---
				N-NH4	-		
				d15N-NO3	-		
				d15N-NH4	-		
				d15N	-		
				DIN	-		
				d13C	-		
				densidade_aparente	-		

Classe	Descrição			Código	Primitiva Geométrica		
Unidade	Descreve as informações relacionadas as unidades utilizadas para medição das amostras.			1.17			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome_unidade	Alfanumerico	20	Nome da unidade utilizada para a medição	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Amostras_solo	Classe principal das informações para solos			1.18			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Conteudo_argila	Real	-	O conteúdo de argila no intervalo de profundidade de solo relacionado em %.	A ser preenchido	---	Nulo	---
Ano_amostra	Inteiro	-	O ano que amostra ou processos de	A ser	---	Nulo	---

			laboratório foram realizados	preenchido			
Periodo_amostra	Alfanumerico	-	Mês ou temporada que a amostra foi realizada	A ser preenchido	---	Nulo	---
Area_plotada	Real	-	Informação sobre o tamanho da área de amostra em metros quadrados	A ser preenchido	---	Nulo	---
Valor_densidade_aparente	Real	-	A densidade aparente do solo no intervalo de profundidade relacionado	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Horizonte	Descreve as informações respeito das profundidades onde foram pegas as amostras			1.19			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
intervalo_inicial	Real	-	Gama inicial do intervalo onde a amostra foi pego, em cm.	A ser preenchido	---	Nulo	---
intervalo_final	Real	-	Gama final do intervalo onde a amostra foi pego, em cm.	A ser preenchido	---	Nulo	---
Profundidade_composta	Booleana	-	Indica se o intervalo de solo selecionado é composto por outros dois intervalos	True	É composto	Nulo	---
				False	Não é composto		

Classe	Descrição			Código	Primitiva Geométrica		
Química_amostra	Descreve as informações respeito das informações químicas que foram pegas			1.20			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia

Al	Real	-	Conteúdo de alumínio na amostra pegada, em cmol dm-3.	A ser preenchido	---	Nulo	---
CEC	Real	-	Capacidade de troca da legenda (Caption Exchange Capacity), em cmol dm-3.	A ser preenchido	---	Nulo	---
Fe	Real	-	O conteúdo de ferro na amostra pegada no intervalo selecionado, em cmol dm-3.	A ser preenchido	---	Nulo	---
Si	Real	-	O conteúdo de silício na amostra pegada no intervalo selecionado, em cmol dm-3.	A ser preenchido	---	Nulo	---
pH_no_H2O	Real	-	pH do solo em água destilada	A ser preenchido	---	Nulo	---
pH_no_CaCl2/KCl	Real	-	pH do solo em solução de sal	A ser preenchido	---	Nulo	---
WFPS	Real	-	Espaço dos poros cheios de água, em %.	A ser preenchido	---	Nulo	---

MODELO DA VEGETAÇÃO

Classe	Descrição			Código	Primitiva Geométrica		
Referências	Referência completa onde a amostra foi pego.			1.01	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Observacoes	Alfanumérico	150	Refere-se a algum comentário extra sob o trabalho	A ser preenchido	---	Nulo	---
Referencias	Alfanumérico	150	Referência completa do trabalho utilizado	A ser preenchido	---	Nulo	---
URL	Alfanumérico	300	URL onde foi encontrada a referência	A ser preenchido	---	Nulo	---
Tipo_referencia	Alfanumérico	80	Tipo de referência	Artigo publicado	Artigos publicados em eventos científicos.	Nulo	---
				Dissertação (mestrado ou doutorado)	Dissertação(M) de tese(D).		
				Capitulo de livro	Capitulo específico dum livro.		
				No campo	Dados obtidos no campo mesmo		

Classe	Descrição			Código	Primitiva Geométrica		
Classificacao_Solo	Classificação do solo de acordo à nomenclatura brasileira, USDA e FAO (Food and Agriculture Organization of the United Nations)			1.02	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Classificacao_brasileira	Alfanumérico	100	Classificação do solo segundo a nomenclatura brasileira	A ser preenchido	---	Nulo	---
USDA_soil_taxonomy	Alfanumérico	100	Classificação do solo segundo a nomenclatura dos Estados Unidos	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Mudancas_Uso_Terra	Classificação do solo de acordo à nomenclatura brasileira, USDA e FAO (Food and Agriculture Organization of the United Nations)			1.03	---		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Tipo_uso_solo_atual_idade	Inteiro	-	Numero de anos que a terra foi convertida.	A ser preenchido	---	Nulo	---
Desde_ano	Tempo	-	Ano em que iniciou o tipo de vegetação real	A ser preenchido	---	Nulo	---
Especies_utilizadas	Alfanumérico	100	Principais espécies no uso da terra	A ser preenchido	---	Nulo	---
Tratamento	Alfanumérico	100	Tipo de tratamento da terra	A ser preenchido	---	Nulo	---
Vegetacao_Predominante	Alfanumérico	100	Vegetação predominante da área	A ser preenchido	---	Nulo	---
Tipo_uso_terra	Alfanumérico	50	Tipo de gestão na área da terra	A ser preenchido	---	Nulo	---
Tipo_uso_solo_atual	Alfanumérico	100	Descreve o tipo de uso do solo real	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Município, Estado	Pontos espaciais referenciais onde a amostra foi tomada			1.04	<input type="text"/>		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Geocodigo	Alfanumérico	7	Código criado pelo IBGE para identificar as unidades Político Administrativas da Divisão Territorial Brasileira e suas subdivisões operacionais (setor censitário), compondo chave única das unidades de coleta, apuração e disseminação de dados estatísticos e territoriais.	A ser preenchido	---	Nulo	---

Nome	Alfanumérico	80	Nome da municipalidade	A ser preenchido	---	Nulo	---
Geometria_aproximada	Boleano	-	Indica que a geometria adquirida é aproximada, em relação à escala prevista para o produto cartográfico.	Sim	---	Não Nulo	---
				Não	---		
Nome_abreviado	Alfanumérico	20	Nome da municipalidade abreviado	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Informacao_amostra	Informações referentes a cada amostra pegada			1.05			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
periodo_amostra	Alfanumérico	20	Informação do mês em que amostra foi pego. Ex. outubro, novembro.	Meses do ano.	---	Nulo	---
qtde_amostra	Inteiro	-	Numero de amostras pegadas.	A ser preenchido	---	Nulo	---
area_mostrada	Real	-	Informação sobre o tamanho da área da amostra em metros quadrados.	-	---	Nulo	---
ano_amostra	Inteiro	-	O ano que a amostra ou procedimento de laboratório foi executado.	A ser preenchido	---	Nulo	---
estacao_amostra	Alfanumérico	30	Estação do ano em que a amostra foi pego.	Seca	---	Nulo	
				Húmida			

Classe	Descrição			Código	Primitiva Geométrica		
Localizacao_geografica	Informações da localização geográficas da amostra.			1.06	★		
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia

local	Alfanumérico	30	O nome do local onde a amostragem foi executada	A ser preenchido	---	Nulo	---
Área_local	Real	-	Área do local em Hectares	A ser preenchido	---	Nulo	---
Coordenadas_obtidas_em	Alfanumérico	10	Local onde as coordenadas foram obtidas	CPM	Ponto central do município	Nulo	---
				SP	Em algum lugar do local		
Decimal_latitude	Alfanumerico	20	Latitude decimal da amostra	A ser preenchido	---	Não Nulo	---
Decimal_Longitude	Alfanumerico	20	Longitude decimal da amostra	A ser preenchido	---	Não Nulo	---
GeometriaAproximada	Booleano	-	Indica que a geometria adquirida é aproximada, em relação à escala prevista para o produto cartográfico.	Sim	---	Não Nulo	---
				Não	---		

Classe	Descrição			Código	Primitiva Geométrica		
Temperatura_anual	Descreve as informações relacionadas à temperatura anual do local estudado.			1.07			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Temperatura_minima	Real	-	Indica a temperatura mínima anual média do ar em graus célsius.	A ser preenchido	---	Nulo	
Temperatura_maxima	Real	-	Indica a temperatura máxima anual média do ar em graus célsius.	A ser preenchido	---	Nulo	---
Temperatura_media	Real	-	Indica a temperatura média anual do ar em graus célsius.	A ser preenchido	---	Nulo	---

Media_historica	Real	5	Temperatura do ar média anual para uma série histórico específica.	A ser preenchido	---	Nulo	---
-----------------	------	---	--	------------------	-----	------	-----

Classe	Descrição			Código	Primitiva Geométrica		
Precipitacao_anual	Descreve as informações relacionadas à precipitação anual do local estudado.			1.08			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nro_meses_estacao_seca	Int	-	Número de meses por ano, que caracteriza a estação seca da região de amostragem.	A ser preenchido	---	Nulo	---
Precipitacao_anual	Real	-	Precipitação anual em milímetros.	A ser preenchido	---	Nulo	---
Media_historica	Real	-	Precipitação média anual, em milímetros.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Caracterizacao_local	Classe que contém informações relativas à precipitação, temperatura, classificação do solo, altitude e bioma.			1.09			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Altitude	Real	-	Altitude respeito do nível do mar.	A ser preenchido	---	Nulo	---
Bioma	Alfanumérica	30	Nome do bioma brasileiro onde a amostra foi pego.	Amazônia	Constituída principalmente por uma floresta tropical	Nulo	---
				Cerrado	Constituído principalmente por savanas.		
				Mata	Constituída principalmente por		

				Atlântica	mata ao longo da costa litorânea que vai do Rio Grande do Norte ao Rio Grande do Sul.		
				Caatinga	Constituída principalmente por savana estépica.		
				Pampas	Constituído principalmente por vegetação campestre.		
				Pantanal	Constituído principalmente por savana estépica alagada em sua maior parte.		
Altitude_obtida_em	Alfanumérica	5	Local onde a altitude foi obtida.	CPM	Ponto central do município.	Nulo	---
				SP	Ponto do local		

Classe	Descrição			Código	Primitiva Geométrica		
composicao	Descreve as informações relacionadas ao Fosforo, carbono, nitrogênio, d15N, d13C, planta_C/N, lignina e lignina/N do local estudado.			1.10			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Intervalo_elemento	Real	-	Varição (mínimo - valor máximo) no fosforo, carbono, nitrogênio, d15N e d13C, planta_C/N, lignina e lignina/N foliar para a comunidade de plantas ou de um grupo de indivíduos da mesma espécie, em %.	A ser preenchido	---	Nulo	---
SE/SD_elemento	Real	-	O valor do erro padrão ou desvio Fosforo, Carbono, Nitrogênio, d15N, d13C, planta_C/N, lignina e lignina/N para a planta, em %.	A ser preenchido	---	Nulo	---
Tipo_DV	Alfanumérico	5	Tipo de dispersão estatístico.	SE	Erro standard	Nulo	---

				SD	Divergência standard		
Nome_elemento	Alfanumerico	20	Nome dos elementos químicos estudados	Fosforo	-	Nulo	---
				Carbono	-		
				Nitrogênio	-		
				d15N	-		
				d13C	-		
				Planta_C/N	-		
				Lignina	-		
				Lignina/N	-		

Classe	Descrição			Código	Primitiva Geométrica		
Descricao_taxonomica	Descreve as características taxonômicas da amostra selecionada num intervalo.			1.11			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
cf	Alfanumérico	5	Incluído "aff."	aff.	Quando o epíteto não é confirmada.	Nulo	---
Epíteto	Alfanumérico	30	O nome específico das espécies.	A ser preenchido	---	Nulo	---

Variedade_subespécies	Alfanumerico.	30	Referem-se ao grau inferior ao da espécie.	A ser preenchido	---	Nulo	---
Nome_comum	Alfanumérico	30	nome regional ou local das espécies vegetais	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Gênero	Descreve nome do gênero das espécies.			1.12			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumerico	-	Nome do gênero das espécies	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Família	Descreve nome do gênero das espécies.			1.13			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumerico	-	Nome da família das espécies	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Autor	Descreve o nome do autor que descreveu a espécie			1.14			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumerico	-	nome do autor que descreveu a espécie	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Reabsorcao	Descreve a porcentagem de reabsorção, é baseada em mudanças na concentração de nitrogênio e fosforo por unidade de massa foliar, pode ser para uma única folha da média, para um indivíduo ou espécie.			1.15			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nome	Alfanumerico	80	nome do elemento medido	Nitrogênio	---	Nulo	---
				Fosforo			
Valor	Real	-	Valor em porcentagem	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Caracterizacao_especies	Descreve as características das espécies da amostra selecionada num intervalo.			1.16			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Leg_X_no_leg	Alfanumérico	30	Se a espécie é uma leguminosa independente se corrige ou não N atmosféricas ou não.	L NL	Espécies de leguminosas Espécies não leguminosas	Nulo	---
Forma_vida	Alfanumérico	30	Tipo de forma de vida.	Tree Shrub Herb Grass epiphyte Hemiepiphyte Liana Macrophyte	Arvore Arbusto Erva Gramínea Epifítico Hemiepífita Cipó Macrófitas	Nulo	---
Potencial_de_fixacao_N	Alfanumérico	5	Refere-se à habilidade de a leguminosa ter rizóbio associação	Yes No	Associação rizóbio Não há associação rizobium.	Nulo	---
Grupo_funcional	Alfanumérico	30	O grupo funcional ao qual ele pertence.	Subcanopy lower canopy upper canopy top canopy young adult pioneer late secondary	subdosseo Inferior no dosseo Superior no dosseo emergente jovem adulto pionero Sucessional secundário	Nulo	---

					tardio		
				early secondary	Sucessional secundário inicial		
Tipo_fenologia	Alfanumerico.	30	Fenologia da folha	deciduous	decíduo	Nulo	---
				semi-deciduous	Semi-decíduo		
				evergreen	Sempre verde		
Fixacao_N2_estimada	Real	-	Descreve uma estimativa quantitativa da fixação do N2 atmosférico realizada pelas leguminosas	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Morfometria	Descreve as informações relacionadas à morfometria do local estudado.			1.17			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Area_foliar	Real	-	Valor médio de área foliar da comunidade vegetal ou dentro de um grupo de indivíduos da mesma espécie ou o referente valor individual de um único indivíduo amostrado.	A ser preenchido	---	Nulo	---

Área_foliar_unidade_usada	Alfanumerico	5	Unidade usada para medir a área de folha.	A ser preenchido	---	Nulo	---
Massa_folha_seca	Real	-	Valor médio de massa seca de folhas para a comunidade de plantas ou dentro de um grupo de indivíduos da mesma espécie ou o referente valor individual de um único indivíduo amostrado.	A ser preenchido	---	Nulo	---
Massa_folha_seca_unidade_usada	Alfanumérico	5	Unidade usada para medir a massa de folha seca.	A ser preenchido	---	Nulo	---
SLA	Real	-	(Specific Leaf Area) Medida da área foliar / massa seca de folhas; pode ser para uma única folha da média para um indivíduo ou espécie.	A ser preenchido	---	Nulo	---
SLA_unidade_usada	Alfanumerico	5	Unidade usada para medir SLA.	A ser preenchido	---	Nulo	---

LMA	Real	-	(Leaf Mass per Area) medida da área foliar / massa seca de folhas; pode ser para uma única folha da média para um indivíduo ou espécie.	A ser preenchido	---	Nulo	---
LMA_unidade_usada	Alfanumerico	5	Unidade usada para medir LMA	A ser preenchida	---	Nulo	---
Espessura_tecido	Real	-	O valor médio de espessura do tecido para a comunidade de plantas ou de um grupo de indivíduos da mesma espécie ou o referente valor individual de um único indivíduo amostrado, em mm.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Serapilheira	Informações da serapilheira no intervalo de solo selecionado.			1.18			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Teor_N	Real	-	Concentração de nitrogênio na serapilheira tomadas no chão, em %.	A ser preenchido	---	Nulo	---
Teor_P	Real	-	Concentração de fosforo na serapilheira tomadas no chão, em %.	A ser preenchido	---	Nulo	---
Taxa_decomposicao_K	Real	-	O valor de K para a comunidade	A ser preenchido	---	Nulo	---
queda_total	Real	-	A quantidade de lixo produzido pela comunidade de plantas em um período de um ano, em kg ha-1yr-1.	A ser preenchido	---	Nulo	---
Queda_folhas	Real	-	A quantidade de serapilheira produzida pela comunidade de plantas em um período de um ano, em kg ha-1yr-1.	A ser preenchido	---	Nulo	---
deposicao_N	Real	-	É a relação de serapilheira produzida pela comunidade de plantas em um período de um ano pela concentração de nitrogênio na serapilheira, em kg ha-1yr-1.	A ser preenchido	---	Nulo	---
deposicao_P	Real	-	É a relação de serapilheira produzida pela comunidade de plantas em um período de um ano pela concentração de fósforo na serapilheira, em kg ha-1yr-1.	A ser preenchido	---	Nulo	---

Classe	Descrição			Código	Primitiva Geométrica		
Vegetacao	Classe principal da vegetação, ela contém alguns elementos não classificados.			1.19			
Atributo	Tipo	Tamanho	Descrição	Domínio	Descrição	Requisito	Fotografia
Nível_estudo	Alfanumérico	10	Descreve o nível de estudo realizado no local.	C	Comunidade	Nulo	---
				GS	Grupo de diferentes espécies. Dados relativos a espécies diferentes de uma mesma comunidade		
				S	Espécies. Dados de diferentes indivíduos de uma espécie específica.		
				I	Nível individual		
Parte_planta	Alfanumérico	20	Descreve o estudo de alguma parte da planta.	Folha	Folhas maduras e saudáveis		
				Folha jovem	Folhas no primeiro estágio de desenvolvimento.		
				Caule	---		
				Ramo	---		
				Fruta	---		
				Flores	---		
				Raiz	---		
				Planta inteira	Significa que a planta inteira foi tomada para amostragem e cada parte da planta foi considerada como uma amostra		
serapilheira							

Anexo “B”: Modelo de dados conceitual e logico para Solos e Vegetação

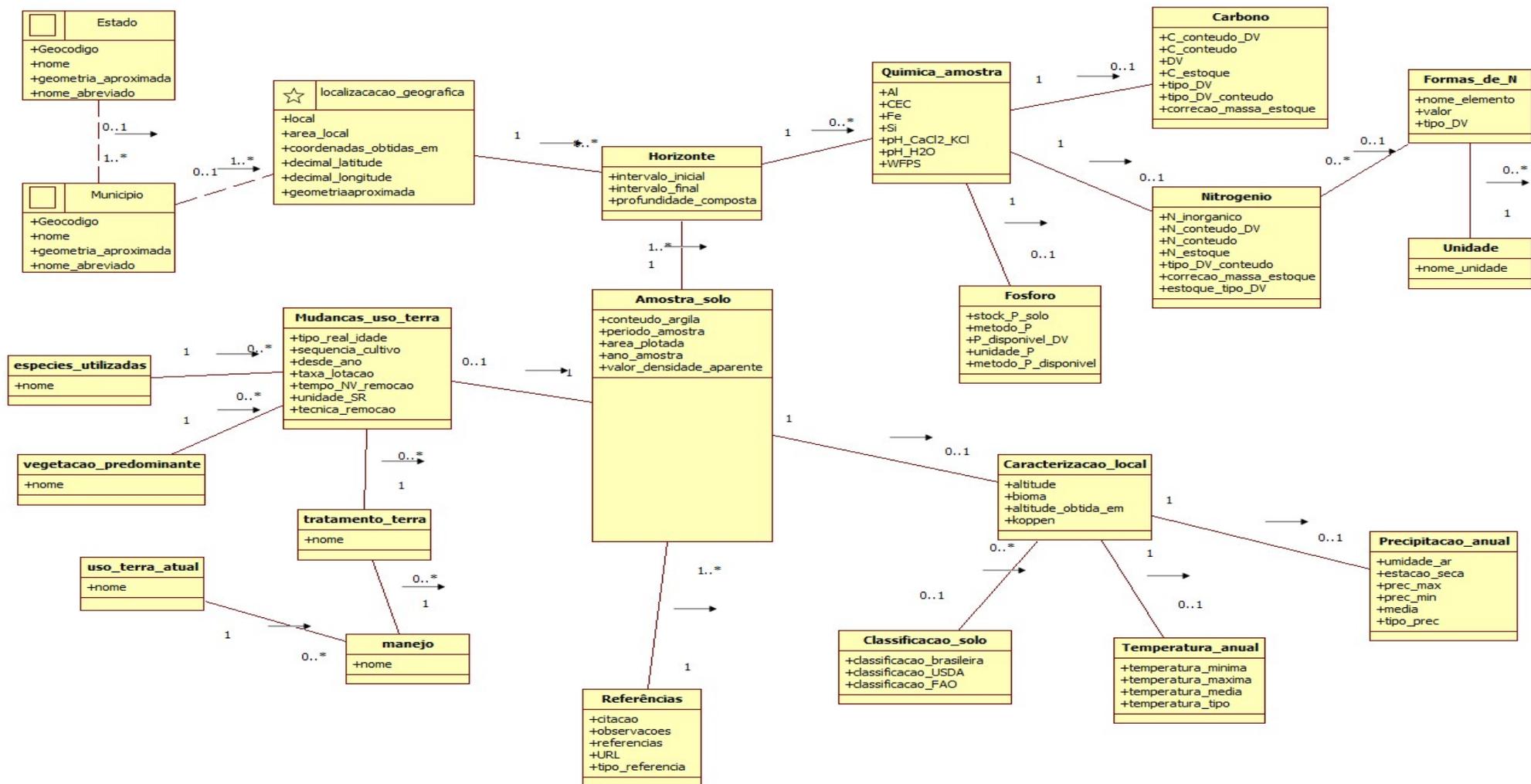


Figura 1 Modelo de dados relacional dos solos.

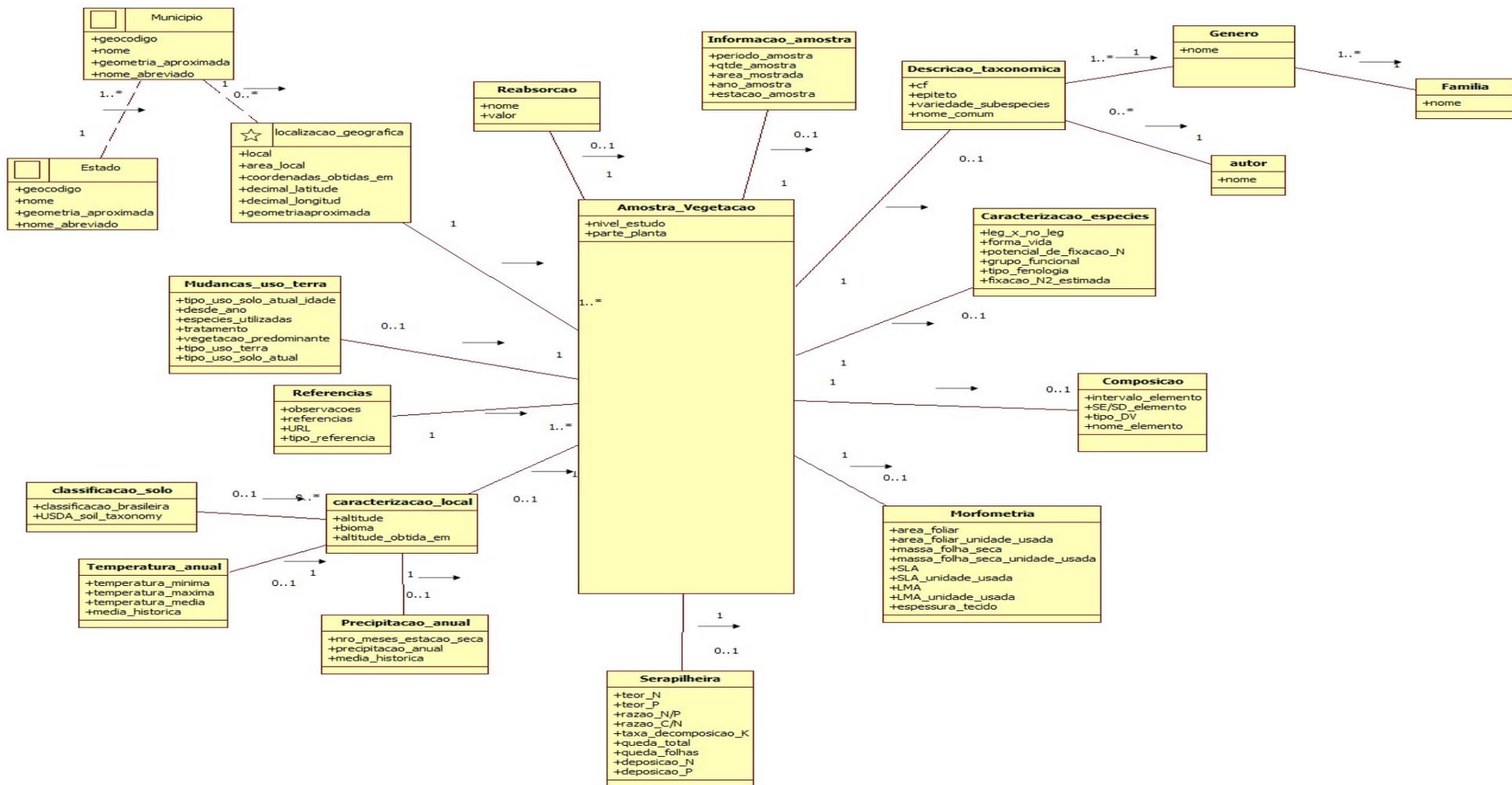
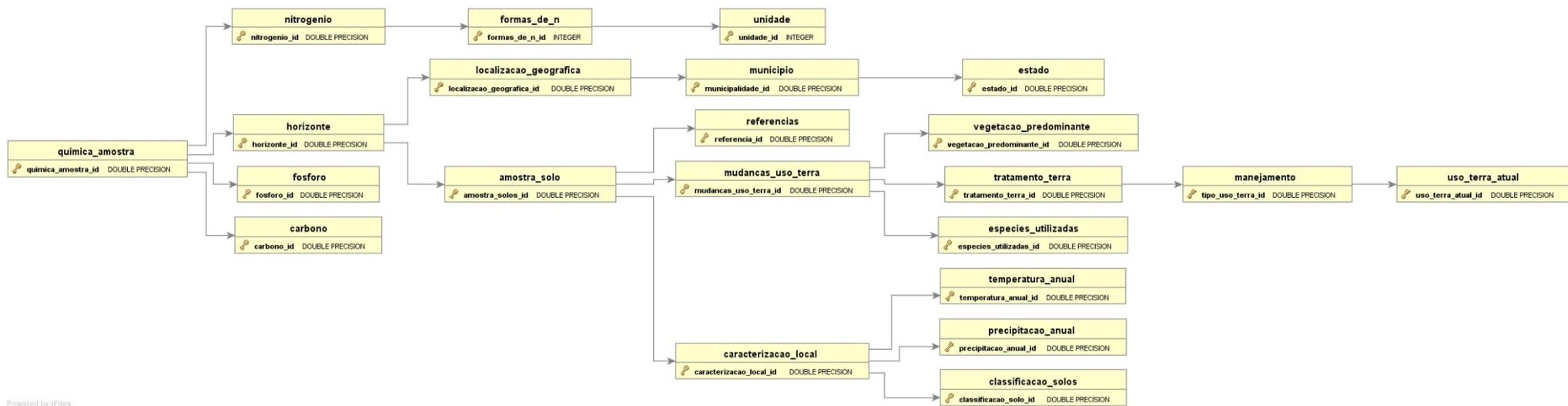
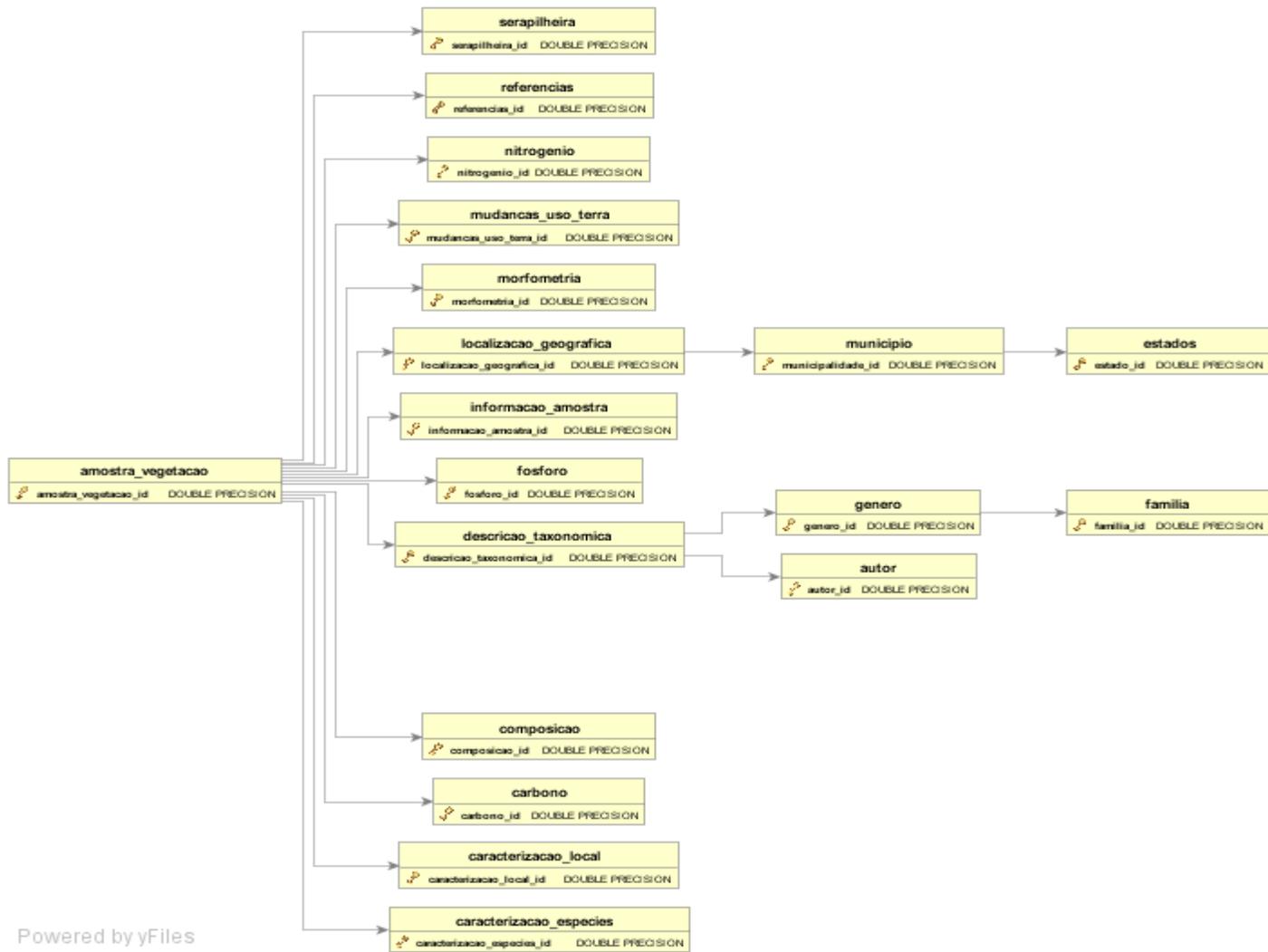


Figura 2 Modelo de dados relacional da vegetação.



Powered by yFiles

Figura 3 Modelo de Dados Logico do Solo.



Powered by yFiles

Figura 4 Modelo de Dados Logico da Vegetação