

Sistema de Informações Geográficas: Uma Abordagem Sobre Interoperabilidade Apoiada em Ontologias e XMI

Alfredo L. P. Manhães
*Programa de Pós-graduação
em Eng. de Computação -
Geomática - Universidade do
Estado do Rio de Janeiro -
UERJ*
alfredomanhaes@yahoo.com

Neide dos Santos
*Programa de Pós-graduação
em Eng. de Computação -
Geomática - Universidade do
Estado do Rio de Janeiro -
UERJ*
neide@ime.uerj.br

Oscar L. R. de Farias
*Programa de Pós-graduação
em Eng. de Computação -
Geomática - Universidade do
Estado do Rio de Janeiro -
UERJ*
fariasol@eng.uerj.br

Resumo

A comunidade que desenvolve sistemas de informação vem ao longo dos anos buscando meios que facilitem a troca de dados e informações entre ambientes heterogêneos, o que deve ser realizado à luz de um poder de expressão universal, favorecendo a interoperabilidade sintática e semântica.

Este trabalho apresenta uma proposta para modelagem de um Sistema de Informações Geográficas (SIG), utilizando uma ontologia para selecionar os elementos do domínio a serem retratados nos mapas e o padrão XML based Metadata Interchange – XMI. Com esta abordagem busca-se atingir a interoperabilidade semântica entre aplicações, mais especificamente entre o ambiente de edição de ontologias e o ambiente de SIG propriamente dito.

1. Introdução

Interoperabilidade, de acordo com Wegner [1], é “a habilidade de dois ou mais componentes de software cooperarem, a despeito de suas diferenças de linguagem, interface e plataforma de execução”. Já Heiler [2] a define como “a habilidade de sistemas distribuídos em larga escala trocarem serviços e dados entre si”. As duas definições abordam aspectos diferentes da interoperabilidade, mas são complementares, já que a primeira trata da interação entre componentes de software e plataformas distintos, ao passo que a segunda se preocupa com os aspectos funcionais, não importando as diferenças estruturais em análise.

Segundo Medeiros e Alencar [3] os estudos nesta área tratam de três tipos de problemas em geral, envolvendo (1) interfaces que permitam o trabalho cooperativo entre

usuários, (2) processos para troca de mensagens e dados entre sistemas e (3) dados para reutilização e intercâmbio. Os mesmos autores mostram que, com base nesses três enfoques, as pesquisas em interoperabilidade apontam para uma redução nos custos em atividades relacionadas à execução de aplicativos, uso de dados, treinamento, implementação e manutenção de aplicações. Aparício et al [4] citando Shet (1998), aponta quatro níveis de heterogeneidade: Sistemico, Estrutural, Sintático e Semântico.

Desta forma, o objetivo do presente trabalho é explorar o tema interoperabilidade e apresentar uma proposta de modelagem para SIG baseada em uma ontologia de domínio que seja interoperável, capaz de nortear a definição de elementos como classes, atributos e relacionamentos entre classes, associados a domínios específicos. Para analisar essa proposta, foi desenvolvido um estudo de caso, cujo domínio é formado por elementos do Zoneamento Municipal proposto pelo Plano Diretor do município de Macaé-RJ. Os elementos do domínio foram utilizados na construção de um protótipo e Sistema de Informações Geográficas (SIG).

Com este propósito, a seção 2, descreve brevemente a base teórica, a saber: Ontologias, Sistemas de Informação Geográfica e *XML based Metadata Interchange – XMI*; a seção 3 apresenta o estudo de caso; e por fim a seção 4 oferece as conclusões do artigo.

2. Base Teórica

2.1. Ontologias

O termo ontologia é definido por Guarino [5] como “uma teoria lógica para relacionar o significado pretendido de um vocabulário formal, isto é, seu comprometimento com uma conceitualização particular

do mundo”. Uma ontologia pode ser entendida [6] como uma descrição explícita de um domínio, ou ainda, um vocabulário comum relacionado a um entendimento compartilhado onde são descritos: (1) Conceitos, (2) Propriedades e atributos do conceito, (3) Restrições às propriedades e atributos e (4) Instâncias.

As ontologias podem ser classificadas conforme mostrado na figura 1. As ontologias de Alto Nível ou de Topo descrevem conceitos gerais, como espaço, tempo, matéria, objeto, evento, ação etc, que são independentes de um problema particular ou domínio. As ontologias de Domínio descrevem um vocabulário relacionado a um domínio genérico, como medicina, direito, cadastro técnico etc. As ontologias de Tarefas descrevem uma tarefa ou uma atividade genérica, como diagnóstico, vendas, etc, pela especialização de termos introduzidos na ontologia de alto nível. As ontologias de Aplicação descrevem conceitos que dependem tanto de um domínio específico como de uma tarefa específica, e que geralmente são uma especialização de ambos.

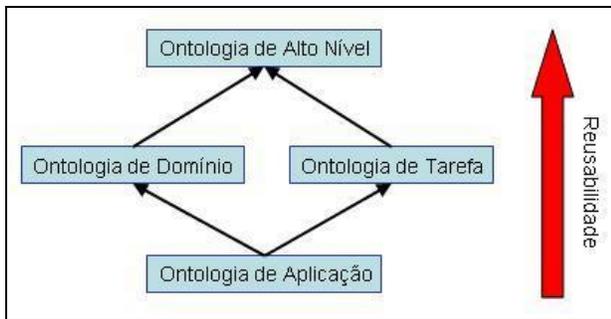


Figura 1 - Tipos de ontologias.

Uma ontologia permite a montagem do vocabulário específico de um domínio, de forma que ele seja modelado em um dicionário de termos, que servirá de base para a especificação dos requisitos de uma aplicação, facilitando sua documentação, manutenção e permitindo a reutilização de conhecimento.

Embora não haja um consenso sobre a melhor metodologia para construção de ontologias, para o desenvolvimento do estudo de caso deste trabalho foi adotada a abordagem de Noy e McGuinness [7], que propõe a elaboração de uma ontologia com as seguintes etapas: (1) determinar o domínio e o escopo da ontologia, (2) investigar o reuso de ontologias existentes, (3) listar termos importantes, (4) definir as classes, (5) identificar a hierarquia de classes e (6) definir propriedades das classes.

Para a edição da ontologia de domínio do estudo de caso foi escolhido o Protégé 3.2.1, desenvolvido pelo Departamento de Informática Médica da Universidade de

Stanford, cujo projeto original tratava da aquisição de conhecimento para a área de Cancerologia. Ele foi desenvolvido em plataforma Java, com código aberto, e deve ser instalado diretamente no computador, distinguindo-o de outros editores que só estão disponíveis em ambiente Web.

2.2. Sistemas de Informação Geográfica

Um Sistema de Informação Geográfica (SIG) é “um sistema para capturar, armazenar, verificar, manipular e visualizar dados georeferenciados” [8], e trata de dados que estão associados às suas respectivas posições no globo terrestre, ou seja, dados que geram informações georeferenciadas. De acordo com Câmara et al. [9], um SIG tem como principais características (1) integrar, numa base de dados única, informações espaciais obtidas a partir de diversas fontes de dados, e (2) oferecer ferramentas computacionais que permitam a combinação dos diversos tipos de informação. De uma forma geral um SIG apresenta os componentes ou subsistemas, representados na figura 2.

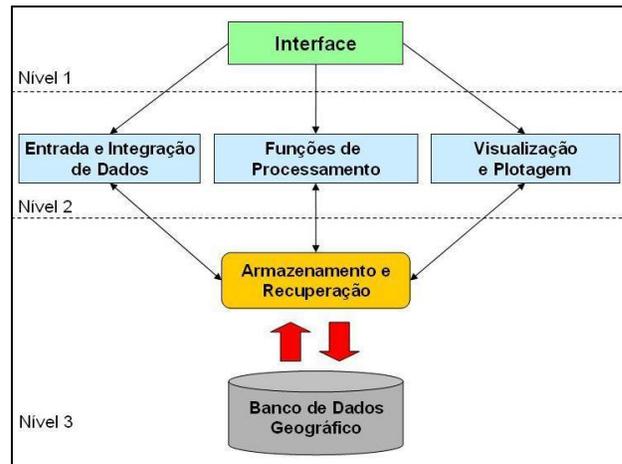


Figura 2 - Componentes de um SIG

A representação das imagens digitais utilizadas no SIG pode ser feita em formato matricial e vetorial. Na forma matricial, também conhecida como raster ou bitmap, o espaço geográfico é representado por uma matriz $P(m,n)$ contendo a informação dos pixels formadores das imagens. A representação vetorial mostra feições da superfície terrestre mapeadas em um plano bi ou tridimensional na forma de pontos, linhas e polígonos, e associados a um sistema de coordenadas para georeferenciar a localização desses elementos geométricos. Cada elemento é uma entidade

independente com propriedades como cor, forma, traçado, tamanho e posição na tela, incluídas na sua definição.

Os SIG fornecem ferramental importante para auxiliar e automatizar a Cartografia, área do conhecimento que compreende um conjunto de atividades científicas, tecnológicas e artísticas, com o objetivo de representar graficamente a superfície terrestre e o universo, na forma de mapas e cartas.

2.3. XML based Metadata Interchange – XMI

Uma linguagem de marcação ou markup language é um agrupamento de símbolos que combina textos ou dados a informações particulares ou propriedades específicas deles, permitindo representar tanto o conteúdo quanto as instruções relativos à sua estrutura. A *eXtensible Markup Language* - XML é uma linguagem de marcação definida pelo *World Wide Web Consortium* - W3C [10]. Em XML, assim como na HTML, a marcação é feita na forma de marcadores ou *tags*, que se destacam dos textos descritores de dados.

O XML based Metadata Interchange – XMI [11] é uma especificação baseada na *Meta Object Facility* – MOF, onde ambos são padrões definidos pelo *Object Management Group* - OMG. MOF é uma tecnologia para definição de metamodelos, oferecendo uma arquitetura para o desenvolvimento de repositórios de metadados que podem ser intercambiados entre aplicações e ferramentas de engenharia de software. O termo metadado (do grego *meta-* + latim *data*) se refere a dados sobre dados, e segundo Marco [12] é definido com *"todo dado ou informação que resulte em conhecimento relacionado aos processos técnicos ou de negócio de uma organização"*.

A especificação XMI se traduz em um padrão de interoperabilidade que favorece a reprodução de classes, atributos e outras características de um domínio sob a ótica da Orientação a Objetos, utilizando XML. Ele permite que ambientes distintos de desenvolvimento de aplicações sejam compatíveis trazendo facilidades, por exemplo, em projetos *Open Source* de sistemas de informação, com alto nível de colaboração entre desenvolvedores.

3. Estudo de Caso: SIG para Zoneamento

O estudo realizado nesse trabalho propõe a modelagem e construção de um SIG para o Zoneamento Urbano do município de Macaé-RJ, utilizando uma ontologia de domínio para selecionar os elementos a serem retratados nos mapas e o padrão XMI para favorecer a interoperabilidade. Com esta abordagem busca-se a interoperabilidade semântica entre aplicações, mais

especificamente entre o ambiente de edição de ontologias (Protégé) e o ambiente de SIG propriamente dito (ArcGIS e ALOV Map). Desta forma, demonstra-se que uma base de conhecimentos pode ser utilizada para a construção de um SIG.

3.1. Etapas de Desenvolvimento do SIG

O desenvolvimento do estudo de caso foi dividido nas etapas mostradas na tabela 1:

TABELA 1 – ETAPAS DO ESTUDO DE CASO

Etapa	Ações
1. Construção da Ontologia de Zoneamento	Estudar o domínio da ontologia.
	Elaborar e validar o Dicionário de Termos (classes) da ontologia.
	Definir as características (atributos) dos termos (classes).
	Construir a ontologia no Protégé.
	Exportar a ontologia para formato XMI.
2. Construção do Diagrama de Classes	Importar o arquivo XMI na ferramenta CASE ArgoUML.
	Ajustar as classes, atributos e relacionamentos no diagrama.
	Gerar o código-fonte das classes (arquivos Java).
3. Construção do protótipo de SIG	Construir tabelas a partir do modelo de classes utilizando o Mapeamento Objeto-Relacional.
	Associar as tabelas geradas com arquivos SHP no ArcGIS.
	Codificar os arquivos XML e HTML para exibição de mapas com o software ALOVMap.
	Exibir mapas no browser.

Não há um consenso sobre a metodologia mais indicada para construção de ontologias. Neste trabalho foi adotada (com adaptações) a abordagem proposta por Noy e McGuinness para a elaboração de uma ontologia, com as seguintes etapas: (1) determinar o domínio e o escopo da ontologia, (2) investigar o reuso de ontologias existentes, (3) listar termos importantes, (4) definir as classes, (5) identificar a hierarquia de classes e (6) definir propriedades das classes.

3.2. Construção da ontologia de Zoneamento

O domínio da ontologia deste trabalho envolve os conceitos definidos na proposta de Zoneamento Urbano do Plano Diretor (PD) do município de Macaé-RJ [14], etapa realizada com especialistas em Urbanismo da Coordenadoria Municipal de Planejamento – COORDEPLAN, consultados na Prefeitura de Macaé e responsáveis pelo PD. As fontes para construção da ontologia envolveram o PD, o Cadastro Imobiliário Municipal, o Cadastro Municipal de Logradouros e a Lei de Uso e Ocupação do Solo do município de Macaé-RJ.

O Dicionário de Termos produzido e validado pela COORDEPLAN, contém os elementos do domínio e seu significado no contexto. Todos esses elementos foram utilizados na construção da ontologia.

A construção da ontologia no Protégé foi realizada, basicamente, a partir da definição das classes (*class*) e suas propriedades (*slot*), além de possíveis instâncias (*instance*), cardinalidades (*cardinality*) e documentação (*documentation*) das classes. Os *slots* do tipo *instance* definem os relacionamentos entre as classes.

A ontologia construída fica disponível para uso em qualquer sistema compatível com os formatos e linguagens do Protégé, como OWL, RDF, XMI, dentre outros. Para atingir os objetivos deste trabalho, optou-se por exportar para o formato XMI, gerando o arquivo onto1.xmi que contém as definições de classes, atributos e relacionamentos do domínio analisado. Essa operação se dá mediante a instalação de um plugin específico que converte a notação da ontologia para UML/XMI. A seguir é apresentado um trecho do arquivo onto1.xmi, com o código XMI relativo ao comentário do slot código_lote, da classe Lote:

```
<UML:Comment xmi.id = 'a380' isSpecification =
    'false' body = 'Código identificador do Lote,
    definido pela Prefeitura.'>
    <UML:Comment.annotatedElement>
    <UML:Attribute xmi.idref = 'a377' />
    </UML:Comment.annotatedElement>
</UML:Comment>
```

A ontologia construída pode ser exportada para outros formatos como RDF, HTML dentre outros, e desta forma ela fica disponível para ser utilizada por outras aplicações e sistemas de informação compatíveis com esses formatos. A figura 3 mostra uma visão parcial da ontologia no software Protégé.

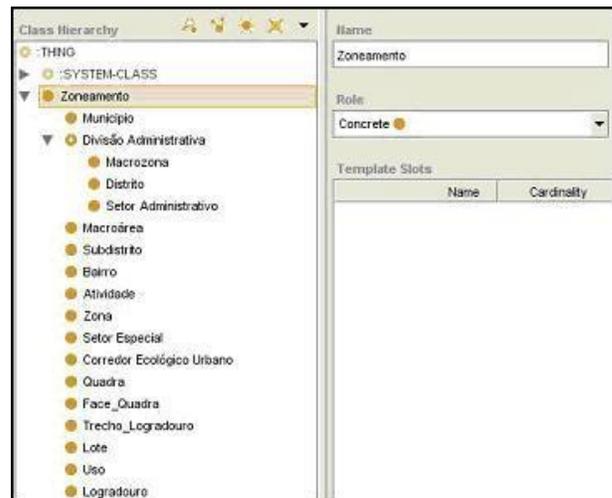


Figura 3 – Visão parcial da ontologia no Protégé.

3.3. Construção do Diagrama de Classes

O arquivo onto1.xmi foi importado pela ferramenta CASE ArgoUML, um software para modelagem de sistemas baseados no paradigma da Orientação a Objetos (OO) e UML. O ArgoUML foi desenvolvido em plataforma Java e é de uso livre.

O ArgoUML permite a geração de código-fonte das classes nas linguagens C#, Java e Php. A título de exemplo é apresentado a seguir o código-fonte Java da classe Atividade:

```
package ProtegeOntology;
class Atividade {
    private final int codigo_atividade;
    private final String descricao_atividade;
}
```

Desta forma, a ontologia construída fica disponível na forma de um pacote de classes, podendo ser utilizada por aplicações que se baseiem no paradigma da OO para o desenvolvimento de sistemas de informação.

Na figura 4 é apresentada a interface do software e o Diagrama de Classes contendo as classes trazidas da ontologia.

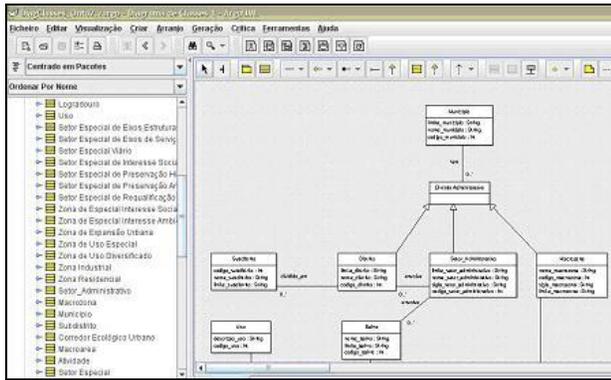


Figura 4 - Classes da ontologia no ArgoUML.

3.4. Construção do Protótipo de SIG

O protótipo de SIG foi construído a partir do mapeamento das classes da ontologia para uma base de dados relacional, resultando em um conjunto de tabelas. Estas tabelas foram unidas por uma operação de junção com imagens vetoriais do município de Macaé, em formato *shapefile*. Essa etapa foi realizada no software ArcGIS. O protótipo para exibição dos mapas foi construído com o software ALOV Map, na versão applet. O ALOV Map é uma aplicação de uso livre, desenvolvida em plataforma Java em 1999 pela empresa russa ALOV, em conjunto com o Departamento de Arqueologia da Universidade de Sidney. O programa é compatível com os formatos SHAPE, MIF, GIF, JPG e MrSID, e permite acesso a diversos bancos de dados relacionais como Access, MySQL e Oracle. A figura 5 representa as tabelas construídas no MS Access.

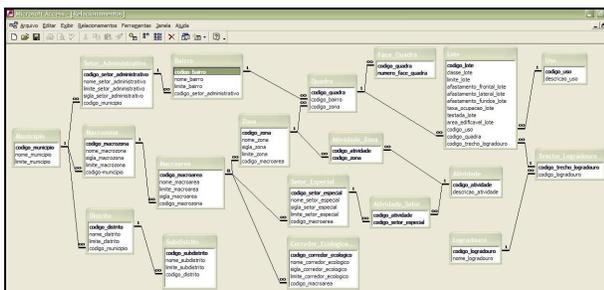


Figura 5 - Tabelas no MS Access.

A visualização de mapas com o ALOV Map necessita de dois arquivos, sendo um em formato HTML e outro com marcação XML. O arquivo sig.html, mostrado a seguir, serve para o carregamento do applet Java, do arquivo camadas.xml, e dimensionamento da área de exibição (*width* e *height*) a serem exibidas no browser.

```
<html>
<head>
<title>Mapas da Ontologia de Zoneamento</title>
</head>
<body>
<applet codebase = "."
code="org.alov.viewer.SarApplet"
archive="alov_applet.jar" width="800" height="600"
align="center">
<param name="pid" value="camadas.xml">
</applet>
</body>
</html>
```

A configuração das feições ou layers a serem exibidas é feita com um arquivo camadas.xml, que contém as marcações referentes ao nome da camada (*layer name*), tipo de arquivo e localização (*dataset url*) e cor em formato RGB para o preenchimento (*symbol fill*) e contorno (*outline*) da imagem. Um trecho desse arquivo é mostrado a seguir, onde são definidas as características de exibição da camada Setores Administrativos.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<project zoomunits="km" mapunits="degrees">
<layer name="Setores Administrativos" visible="yes"
showlegend="yes">
<dataset url="Setores.shp"/>
<renderer>
<symbol fill="190:055:255" outline="38:115:0" />
</renderer>
</layer>
```

3.5. Resultados Obtidos

A exibição dos mapas em ambiente Web foi realizada com o carregamento do arquivo sig.html no browser Firefox (figura 6) onde é mostrada a imagem vetorial contendo os Bairros e Distritos do município de Macaé.

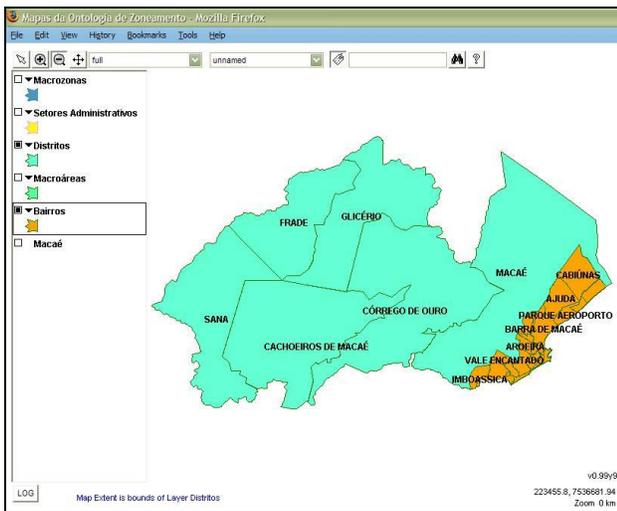


Figura 6 - Visão de Distritos e Bairros.

A figura 7 apresenta um detalhe (zoom) da região que compreende a Macrozona de Ambiente Urbano (MAU), a Macroárea de Uso Sustentável (MUS), os Setores Administrativos Marrom e Vinho, e os bairros Lagomar, Ajuda, São José do Barreto e Parque Aeroporto.

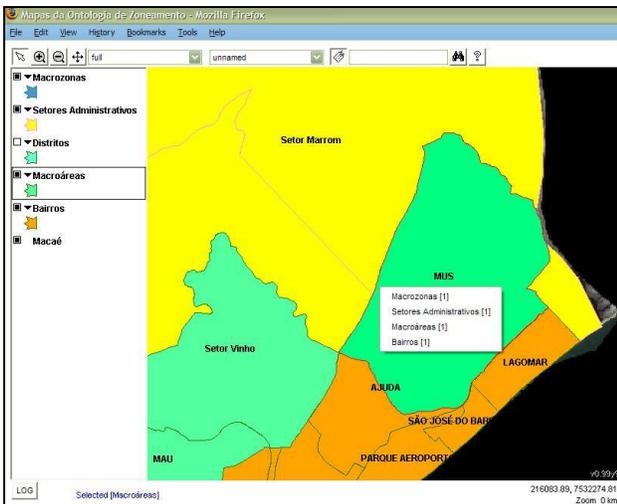


Figura 7 – Superposição de camadas.

A ferramenta de seleção, acionada com o botão em forma de ponteiro na barra de ferramentas do ALOV Map, ao ser clicada em uma parte do mapa, exibe um retângulo branco que mostra as camadas superpostas e o número de linhas de cada tabela associada às camadas visualizadas. A tabela referente à camada Macroárea de Uso Sustentável é mostrada na figura 8, onde podem ser

observadas as colunas provenientes da ontologia de domínio.

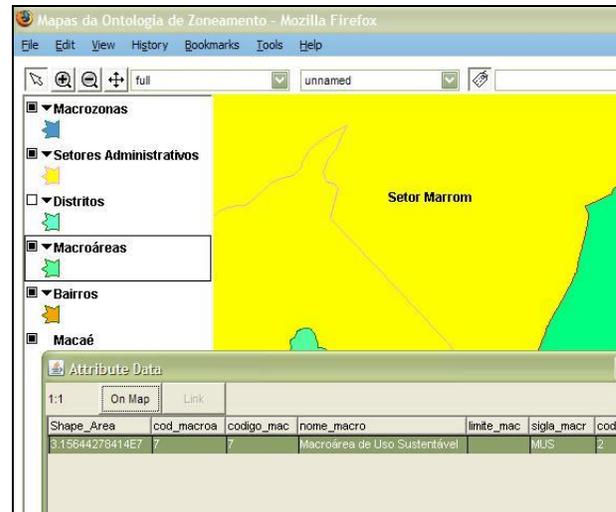


Figura 8 – Seleção da camada Macroáreas.

O SIG construído permite o acesso a todas as informações existentes na ontologia de domínio e favorece sua visualização, tanto na forma de mapas com temas específicos, quanto na forma de tabelas de dados descritivos.

4. Conclusões

A solução apresentada aborda diferentes aspectos da interoperabilidade. A ontologia construída no Protégé tem uma estrutura que segue as recomendações do W3C. Desta forma, o conhecimento do domínio Zoneamento, objeto de estudo deste trabalho, pode ser disseminado e utilizado em diferentes sistemas de informações que sigam a padronização do W3C, adotada no Protégé.

Outro aspecto importante está relacionado à utilização de XMI, que é um padrão especificado pela OMG para dar suporte à modelagem e intercâmbio de metadados entre diferentes ambientes de desenvolvimento de software, e favorecer a interoperabilidade. O estudo propõe um acoplamento entre a ontologia e o SIG utilizando o XMI para favorecer a integração entre esses ambientes, permitindo que o SIG reproduza as informações do domínio em questão.

As classes oriundas da ontologia podem ser reutilizadas em outras aplicações que sejam compatíveis com os formatos *Ontology Web Language* (OWL), *Resource Description Framework* (RDF) e outros utilizados no Protégé. O reuso também pode ser feito a partir do formato XMI, sendo necessário que a ferramenta

CASE que irá manipulá-lo seja compatível, o que já ocorre com ferramentas para modelagem como o Poseidon, ARGO UML, Rational Rose e outros.

As classes definidas na ontologia também podem ser reutilizadas na construção de bancos de dados. Este processo pode ser realizado manualmente ou com a utilização de ferramentas CASE que apresentem esta funcionalidade, e é aplicável a diversos bancos de dados existentes no mercado. Deve-se observar que existem diversas ofertas de software para SIG compatíveis com os bancos de dados mais comuns, o que propicia a reutilização das classes da ontologia para a produção de informação geográfica.

5. Referências

- [1] Wegner, Peter. Interoperability. ACM Computing Surveys, 28(1):285-287, March, 1996.
- [2] Heiler, Sandra. Semantic interoperability. ACM Computing Surveys, 27(2):271-273, June, 1995.
- [3] Medeiros, C. M. B. ; Alencar, A. . Qualidade de dados e interoperabilidade em SIG. In: GeoInfo 1999, 1999, Campinas. Anais I GeoInfo, 1999. p. 12-19. Disponível em <http://www.geoinfo.info/portuguese/geoinfo1999/papers/CBauzer.pdf>
- [4] Aparício, A. S. et al. Integration of Heterogeneous Databases and Ontologies. Universidade do Estado do Rio de Janeiro, 2006. Disponível em http://www306.ibm.com/ebusiness/br/campaign/rational/soa_jornada.shtml.
- [5] Guarino, N. Understanding, Building, and Using Ontologies. Padova: LADSEB-CNR, National Research Council, 1996. Disponível em <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/guarino.html>.
- [6] Parreiras, Fernando Silva. Introdução à construção de ontologias. In: Simpósio Mineiro de Sistemas de Informação, 1, 2004, Belo Horizonte. Disponível em <http://www.fernando.parreiras.nom.br/palestras/engeonto.pdf>.
- [7] Noy, N. F; McGuinness, D. L. Ontology Development 101: A Guide to Creating your First Ontology. Stanford Medical Informatics Technical, 2001. Disponível em http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.
- [8] Burrough, Peter A; McDonnell, Rachel A. Principles of Geographical Information Systems. New York: Oxford, 1998.
- [9] Câmara, Gilberto et al. Anatomia de Sistemas de Informação Geográfica. Campinas: Instituto de Computação, UNICAMP, 1996. Disponível em <http://www.dpi.inpe.br/geopro/livros/anatomia.pdf>.
- [10] W3C. Extensible Markup Language – XML. W3C, 2003. Disponível em <http://www.w3.org/XML/>
- [11] OMG. MOF 2.0 / XMI Mapping Specification, v2.1. 2005. Disponível em <http://www.omg.org/technology/documents/formal/xmi.htm>.
- [12] Marco, David. Building and Managing the Metadata Repository: A Full Lifecycle Guide. New York: Wiley & Sons, 2000.