

# From Geo-Pragmatics to Derivation Ontologies: New Directions for the GeoSpatial Semantic Web

Frederico Fonseca

College of Information Sciences and Technology The Pennsylvania State University

Andrea Rodriguez

Center for Web Research University of Chile and Department of Computer Science

University of Concepcion

As entidades geográficas não estão somente localizadas no espaço. Elas herdam características estruturais do espaço, tais como a mereologia, topologia e propriedades geométricas. Estas características implicam em que a *geo-ontologia* difere de outras ontologias, e incorpora conceitos relativos a topologia e parte-todo característicos dos dados geográficos. Os objetos geográficos podem ser conectados ou contínuos, espalhados ou separados, fechados ou abertos, sendo geralmente complexos formados por várias partes constituintes. Além da dimensão espacial a dimensão temporal também está associadas aos objetos geográficos.

Nesta edição da Transactions on GIS, são abordadas dimensões relativas ao uso da informação geográfica ainda não tratadas nas pesquisas em Web Semântica:

Profissional – informação estruturada e organizada em bancos de dados disponíveis através de portais corporativos.

Amador – informação não estruturada, organizada informalmente nas páginas Web

Científico – informação geográfica organizadas em artigos científicos, modelos e teorias.

## Geo-Pragmatics for the Geospatial Semantic Web

Boyan Brodaric - Geological Survey of Canada Ottawa, Canada

Aborda aspectos do uso científico da informação geográfica.

Entidade Geoespacial – abstrações referindo-se a entidades do tipo região, superfície e relações próximo, distante, ao lado, sobreposto.

Objeto Geográfico – abstrações referindo-se a formações geológicas, construções, montanhas relações tais como “estratigraficamente sobre”, qualificadores tais como altura, grossura e processos nos quais participam, tais como, *falhas (tectônicas)*, *terremotos*, *tsunamis*.

Geocientífico – abstrações descobertas por meio das geociências (ciência dos solos, geologia, geomorfologia, ecologia, hidrologia).

Conceito – engloba a faixa completa de abstrações como ilustradas acima e que engloba no mínimo a definição lógica de cada uma das abstrações.

Conceito científico – interpretado como artefato resultante da intersecção de entidades mentais com a natureza baseado nas relações entre elas.

A aplicação do conhecimento científico para obtenção de classificações e propriedades sobre um determinado objeto geográfico, pode resultar em resultados distintos dependendo do modo como o conhecimento foi utilizado (suposições iniciais e opções escolhidas). Neste contexto é difícil atribuir um grau de “verdade” aos resultados obtidos. Esta situação pode ser um fator limitante no uso de infraestruturas disponibilizando informações espaciais, uma vez que o registro do processo utilizado na obtenção dos dados não é armazenado.

A *Pragmatic Web* visa resolver este problema, onde conceitos são definidos e descrições de padrões

de uso dos conceitos são apresentadas servindo de guia para sua aplicação em contextos reais. A evolução da Web Semântica em direção da *Pragmatic Web*, envolve a adoção de ontologias acompanhadas de uma descrição de seu contexto de uso.

## **A Rule-Based Strategy for the Semantic Annotation of Geodata**

Eva Klien - Institute for Geoinformatics - University of Muenster

Neste artigo é proposta uma estratégia baseada em regras para anotação semântica em dados geoespaciais combinando a Web Semântica e Serviços Web Geoespaciais. As regras automatizam parcialmente o processo de anotação definindo condições para identificação de conceitos espaciais.

### **1 Introdução**

O aumento crescente da disponibilização de informação geográfica por SDI's (*Spatial Data Infrastructures*) gera duas implicações quanto comparado a sistemas GIS monolíticos:

- Um número crescente de usuários com diferentes perfis (experiência, necessidades, etc...) tem acesso a informação geoespacial.
- Os dados geográficos gerados no contexto de uma organização, agora possuem uma comunidade de usuários ampla e heterogênea.

Os padrões sintáticos definidos pela OGC não são suficientes para tornar eficientes os processos de busca e recuperação de fontes de informação geográfica, que atendam as necessidades dos usuários. A utilização de metadados para catalogação de recursos de informação geográficos não levam em consideração as diferentes conceitualizações existentes, provocando uma heterogeneidade semântica na busca e recuperação de informação geográfica.

O problema atualmente existente nestes processos ocorre porque não existe uma representação formal e explícita da semântica possibilitando uma interoperabilidade semântica. A abordagem da Web Semântica, com a especificação de linguagens tais como a OWL (*Web Ontology Language*) e SWRL (*Semantic Web Rule Language*), é complementar a abordagem sintática definida pela OGC, e fornece um caminho para solução da interoperabilidade semântica.

A esta integração se dá o nome de *Geospatial Semantic Web* (Egenhofer 2002, Fonseca and Sheth 2002, Arpinar et al. 2004, Kolas et al. 2005).

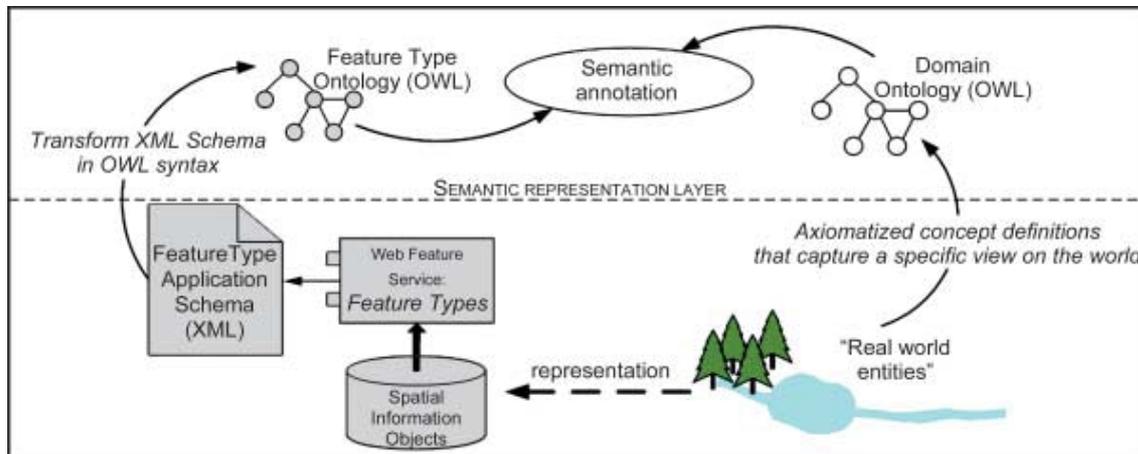
Anotação semântica neste trabalho é entendido como a explicitação das relações entre os esquemas de dados e um domínio ontológico, obtida pelo mapeamento de elementos no esquema em elementos da ontologia. Klien and Lutz (2005), apresentam uma abordagem automatizada para realização da anotação semântica baseada nas relações espaciais. Neste artigo além das relações espaciais é considerado também as propriedades espaciais na detecção de instâncias implícitas no nível da informação.

### **2 General Framework for the Semantic Annotation of Geodata in Web Environments**

As Ontologias definem formalmente conceitos de um domínio e termos para denotá-los. Ontologias podem ser utilizadas para explicitar a semântica da informação geoespacial e desta forma habilitar o casamento semântico automático. Porém a sua adoção em larga escala pela comunidade de usuários de dados geoespaciais depende de um suporte de ferramentas e métodos que facilitem a sua utilização na geração de anotação semântica formal. Neste artigo é apresentado um método para geração automatizada de anotação semântica, que implementa parte das idéias do *semantic reference systems* (Kuhn, 2003).

A Figura 1 ilustra o contexto de aplicação da anotação semântica sobre dados geoespaciais em ambiente Web. As representações digitais das entidades do mundo real receberão a anotação semântica. Estas entidades são representadas como feições codificadas usando GML (OGC, 2007) e são disponibilizadas por meio de WFS (*Web Feature Services*) (OGC, 2007). A única informação explícita disponível sobre estas representações das entidades são os esquemas de aplicação obtidos

do WFS através de requisições do tipo *DescribeFeatureType*. A fim de realizar a anotação semântica (parte superior da figura 1) é necessário definir uma Ontologia para o Domínio Geoespacial, capaz de capturar visões de mundo específicas na forma de axiomas e regras. É necessário também definir uma estratégia para geração de anotação semântica para as feições servidas pelo WFS.



As diferentes visões de mundo dos diferentes grupos de usuários determinam diferentes conceitualizações para descrição do mundo real. Na comunidade de informação geográfica, isto se reflete em diferentes modelos descrevendo a realidade. Uma ontologia pode ser definida para uma comunidade específica, definindo conceitos específicos e um contexto para sua utilização. Segundo Guarino (1998), “uma ontologia é um artefato de engenharia formado por um vocabulário utilizado na descrição de uma certa realidade mais um conjunto de suposições explícitas sobre o significado das palavras neste vocabulário.”

O significado de um termo pode ser classificado em dois tipos de acordo com os filósofos: significado intencional e significado extensional. O significado intencional é definido pelo conceito associado ao termo. O significado extensional de um termo é definido pelo conjunto de entidades reais a que o termo se refere. Geralmente é inviável enumerar todas as entidades a que um termo se refere e, desta forma, o significado intencional é mais utilizado. O significado intencional é construído com base na abstração de entidades do mundo real (significado extensional), através da descrição das propriedades de instâncias típicas destas entidades e de suas relações espaciais com outras entidades. Ontologias do domínio geoespacial tratam do significado intencional dos termos. Dois conceitos podem ser mapeados para um mesmo significado extensional, porém isto não implica que os mesmos tenham um mesmo significado. Por exemplo, uma estrada possui significado diferente para um engenheiro civil e um ecologista.

O significado intencional baseado na descrição de propriedades e relações espaciais de entidades típicas associadas a um conceito pode ser utilizado como um critério para seleção de entidades que formariam o seu significado extensional. A estratégia de anotação semântica de dados geoespaciais esta baseada nesta hipótese.

## 2.2 The Semantic Annotation Task

O significado intencional é construído com base na definição de propriedades das entidades e de suas e relações. Para atribuição de uma anotação semântica sobre um esquema se dá em duas etapas:

- Geração de OWL do Esquema – nesta etapa o esquema XML de uma feição é modificado sintaticamente para constituir uma descrição em OWL. Esta descrição é necessária pois o mapeamento só poderá ser realizado sobre ontologias expressas em uma mesma linguagem. A “ontologia de feições” gerada neste etapa não define uma semântica, mas somente uma taxonomia de feições.

- O segundo passo uma Ontologia de domínio é “executada” sobre um conjunto de entidades, visando identificar quais delas se enquadram como exemplares de um determinado conceito. Nesta verificação pode ser necessária a execução de métodos de análise espacial para identificar se determinadas propriedades ou relações são satisfeitas pelas entidades. A Ontologia de domínio é elaborada por especialistas de um determinado domínio. O processo de anotação semântica consiste na realização de um mapeamento de elementos da “ontologia de feições” na “ontologia de domínio”.

## 2.3 Background on Ontology Mapping

Assumindo que uma ontologia seja representada pela expressão:

$$O = (S, A)$$

onde  $O$  é a ontologia,  $S$  é a assinatura ontológica (descrevendo o vocabulário) e  $A$  um conjunto de axiomas ontológicos definindo restrições sobre significado intencionado do vocabulário. Um mapeamento ontológico total entre duas ontologias  $O_1$  e  $O_2$  é definido (Kalfoglou, 2005) como sendo um mapeamento entre os termos de  $O_1$  em  $O_2$  de forma que:

$\int S_1 \rightarrow S_2$  (mapeamento entre as assinaturas) de modo que  $A_2 \models \int (A_1)$ , isto é, toda interpretação que satisfaça os axiomas em  $O_2$  também satisfaz o axiomas traduzidos de  $O_1$ .

De maneira informal o mapeamento seria um mecanismo de substituição dos termos de uma ontologia por outra mantendo a consistência dos significados.

Na estratégia apresentada neste artigo, a “ontologia de feições” denotada por  $O_{\text{fio}} = (S_{\text{fio}}, ?)$  não expressa restrições explícitas sobre os elementos de seu vocabulário. A ontologia de domínio denotada por  $O_{\text{do}} = (S_{\text{do}}, A_{\text{do}})$  possui um conjunto de axiomas definindo a semântica intencionada dos elementos do vocabulário. O mapeamento de  $O_{\text{fio}}$  em  $O_{\text{do}}$  visa atribuir uma semântica explícita a  $S_{\text{fio}}$ . A ontologia de feições possui um significado estendido representado pelo conjunto de instâncias de feições em uma base de dados.

## 2.4 Strategy for Automatic Support during the Mapping

O mapeamento consiste em confrontar as entidades do banco de dados com as propriedades e relações explicitados nos axiomas da ontologia de domínio, determinando quais delas são exemplares de um determinado conceito. Nesta confrontação podem ser utilizados procedimentos de análise espacial para verificação de algumas destas propriedades.

## 3 Semantic Web Technology

**Web Ontology Language (OWL).**

**Semantic Web Rule Language (SWRL).**

## 4 Case Study for Semantic Annotation

### 4.1 Formalization of the Concept LOWLAND in the Domain Ontology

### 4.2 Transforming the Application Schema into OWL Syntax

### 4.3 Semantic Annotation Process

#### 4.3.1 Detecting string-based similarity

### **4.3.2 Evaluating conceptual similarity by analysing instance data**

### **4.3.3 Refining semantic annotations**

*Mereologia – Estudo lógico-matemático das relações entre as partes e o todo e das relações entre as partes interiores de um todo.*

**Referências**

Kalfoglou, Y., Hu, B., Reynolds, D. and Shadbolt, *Semantic Integration Technologies Survey*. (2005) – Disponível em <<http://eprints.ecs.soton.ac.uk>> acesso em 12/2007.