Chapter XXI
Geospatial and Temporal Semantic Analytics

Matthew Perry
University of Georgia, USA

Amit Sheth
University of Georgia, USA

Ismailcem Budak Arpinar
University of Georgia, USA

Farshad Hakimpour
University of Georgia, USA

ABSTRACT

The amount of digital data available to researchers and knowledge workers has grown tremendously in recent years. This is especially true in the geography domain. As the amount of data grows, problems of data relevance and information overload become more severe. The use of **semantics** *has been proposed to combat these problems (Berners-Lee et al., 2001; Egenhofer, 2002). Semantics refer to the meaning of data rather than its syntax or structure. Systems which can understand and process data at a semantic level can achieve a higher level of automation, integration, and interoperability. Applications generally use semantic technology for three basic purposes: (1) semantic integration, (2) semantic search and contextual browsing, and (3) semantic analytics and* **knowledge discovery** *(Sheth & Ramakrishnan, 2003).*

INTRODUCTION

This chapter focuses on semantic analytics and knowledge discovery in the geographic information science domain. Semantic analytics applications provide capabilities for analyzing relationships and patterns in semantic **metadata**. So far, research in this area has concentrated on thematic relationships between entities (e.g., the fact that two glycopeptides participated in the same biological process). However, for many domains and applications, spatial and temporal relationships cannot be overlooked. Next generation geoinformatics applications that can successfully combine knowledge of real-world entities and relationships with knowledge of their interactions in space and time will have huge potential in areas such as national security and emergency response. The remainder of this chapter reviews background concepts from the Semantic **Web** community and describes state-of-the-art work in semantic analytics and discovery in the purely thematic dimension. It then discusses our ongoing work in realizing semantic analytics and discovery in all three dimensions of information: thematic, spatial, and temporal.

BACKGROUND

In preparation for our discussion of geospatial and temporal semantic analytics, we first review basic concepts of ontologies, for the Semantic Web and thematic analytics.

Ontology

Ontologies are central to realizing semantic applications as they provide a concrete way to specify the semantics of an application domain. Ontology is classically defined as "a specification of a conceptualization" (Gruber, 1993). We can think of an ontology as consisting of two parts: a schema and instance data. The schema models a domain by defining class types (e.g., *University*, *City*) and relationship types (e.g., *located_in*). The schema is populated with instances of classes and relationships (e.g., *The University of Georgia located_in Athens*) to create facts representing knowledge of the domain. A number of ontologies describing thematic aspects of data have been developed at the Large Scale Distributed Information Systems (LSDIS) lab. Some recent examples include GlycO and ProPreO in the Bioinformatics domain (Sahoo et al., 2006) and more general-purpose ontologies such as the Semantic Web Evaluation Ontology (SWETO) (Aleman-Meza et al., 2004).

There has been significant work regarding the use of geospatial ontologies in geographic information science. Ontologies in geographic information systems (GIS) are seen as a vehicle to facilitate interoperability and to limit data integration problems both from different systems and between people and systems (Agarwal, 2005). Fonseca et al. (2002) present an architecture for an ontology-driven GIS in which ontologies describe the semantics of geographic data and act as a system integrator independent of the data model used (e.g., object vs. field). Kuhn (2001) claims that, for maximum usefulness, geo-ontologies should be designed with a focus on human activities in geographic space and thus present a method for constructing domain ontologies based on the text analysis of domain documents (e.g., German traffic code text for the car navigation domain). Kuhn and Raubal (2003) also introduce the concept of semantic reference systems, of which ontologies are a component, as a means to describe the same geographic information from varying perspectives. This includes notions of semantic transformation and projection of ontologies. These operations could potentially be used to present geographic information from different scales and granularities. Frank (2003) goes a step beyond purely spatial ontologies and argues for the inclusion of the temporal dimension by de-