Chapter 6

Using Distributed Semantic Catalogs for Information Discovery on Spatial Data Infrastructures

Fabio Gomes de Andrade
Federal Institute of Education, Science and Technology of Paraiba, Brazil

Fábio Luiz Leite Jr
Paraiba State University, Brazil

Cláudio de Souza Baptista
Federal University of Campina Grande, Brazil

ABSTRACT

Spatial data sharing among both private and public organizations is an old issue. Spatial Data Infrastructures (SDIs) have been proposed to solve data integration and discovery problems. Nonetheless, they can only partially work out these problems, as most of their catalog services are built on keyword based search. This approach always results in low precision on searching results. One of the greatest challenges in spatial integration is to provide the semantics for underlying data and services. This chapter describes a distributed catalog service that uses ontologies to describe underlying information so as to improve searching precision. The proposed catalog service can rewrite and propagate queries to other distributed catalogs to cooperate and exchange information directly without the traditional problems imposed by a centralized architecture.

1. INTRODUCTION

Spatial Data Infrastructures (SDIs) (Williamson, Rajabifard & Feeney, 2003) have become increasingly popular in geospatial communities. These SDIs facilitates the access to and the sharing of geospatial data by providing retrieval mechanisms to integrate underlying data sets. As a result, several initiatives, at country and continental levels, were taken for the development of SDIs,

including the USA NSDI ("National Spatial Data Infrastructure," 2010), the European INSPIRE ("Infrastructure for Spatial Information in the European Community," 2010), the Australian ASDI ("Australian Spatial Data Infrastructure," 2010) and GSDI ("Global Spatial Data Infrastructure," 2010). Some countries have been considering the implementation of nationwide SDIs, as it is the case of Brazil ("Infraestrutura Nacional de Dados Espaciais," 2010).

To design and develop a CDI one has to tackle several problems at the same time. This is due to the complexity inherent to spatial data. Currently, there exist standards for enhancing interoperability in SDI, particularly those provided by the Open Geospatial Consortium (OGC), such as the Web Map Service (WMS) ("Web Map Service Interface," 2004), the Web Feature Service (WFS) ("Web Features Service implementation specification," 2008) and the Web Coverage Service (WCS) ("Web Coverage Service implementation specification," 2008); helps to reduce the complexity regarding access to spatial data. However, many problems still call for effective and efficient information discovery.

Catalog services have been proposed as a way to deal with information discovery in SDI. These catalogs are used for storing and managing information available for both data and services provided by federation SDIs. A client may query these catalog services to discover where geospatial data or services are found. Generally, the information discovery process is based on keywords which relate to the syntax and structure of data sets; however, this process fails to explore the semantics inherent to either the query context or the application. This reduces the precision and the recall of the proposed overall SDI federation. Another important problem related to current SDI architectures is the centralization of catalog services. This centralization also entails some classical distributed system problems, including the existence of a single point of failure, a bottleneck for queries which reduces scalability, and the

complexity of maintaining an updated catalog in relation to dynamic SDI data sources.

This chapter presents the catalog service of the Semantic Web Services - GIS (SWS-GIS) (Leite, Baptista, Silva & Silva, 2007). The SWS-GIS is an architecture for information sharing among members of an SDI federation. The main contributions of this chapter are:

- the proposal of a semantic catalog service that uses inference to improve information discovery and facilitates spatial data integration among heterogeneous data sources; and
- the truly distributed catalog in which queries may be posed from any SDI client without the need of a central catalog. In our approach, each SDI has its own catalog. When a particular query is submitted to any SDI catalog, it is rewritten and propagated to other catalogs.

The remainder of the chapter is organized as follows: Section 2 discusses related works on information discovery in SDIs; Section 3 gives us a background on geospatial semantic web and spatial data infrastructures, focusing on the problems related to spatial data discovery; Section 4 presents the example used to demonstrate our approach; Section 5 highlights the proposed catalog service for the discovery of spatial data in SDIs; and finally, Section 6 presents the conclusions and future directions of this research.

2. RELATED WORKS

Many relevant research works have been proposed for the solution of spatial data discovery problems on SDIs via semantic web concepts. Nevertheless, most of these solutions focus on the use of semantics to get data from a single infrastructure. Such solutions do not consider the possibility of forwarding queries to other SDIs if

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/using-distributed-semantic-catalogs-information/51485

Related Content

The Syntactic and the Semantic Web
J. Cardoso (2007). *Semantic Web Services: Theory, Tools and Applications (pp. 1-23).*
www.irma-international.org/chapter/syntactic-semantic-web/28877

Data Leakage and Privacy Concerns in Public Bug Bounty Platforms
Muhammad Hamad, Altaf Hussainand Majida Khan Tareen (2023). *Protecting User Privacy in Web Search Utilization (pp. 139-159).*
www.irma-international.org/chapter/data-leakage-and-privacy-concerns-in-public-bug-bounty-platforms/322589

Integrating Community Interest and Neighbor Semantic for Microblog Recommendation
Mingxin Ganand Xiongtao Zhang (2021). *International Journal of Web Services Research (pp. 54-75).*
www.irma-international.org/article/integrating-community-interest-and-neighbor-semantic-for-microblog-recommendation/277064

Modern Diffusion of Products with Complex Network Models
Atsushi Tanaka (2011). *E-Activity and Intelligent Web Construction: Effects of Social Design (pp. 119-133).*
www.irma-international.org/chapter/modern-diffusion-products-complex-network/53279

A Novel Completion-Time-Minimization Scheduling Approach of Scientific Workflows Over Heterogeneous Cloud Computing Systems
S. Sabahat H. Bukhariand Yunni Xia (2019). *International Journal of Web Services Research (pp. 1-20).*
www.irma-international.org/article/a-novel-completion-time-minimization-scheduling-approach-of-scientific-workflows-over-heterogeneous-cloud-computing-systems/237998