Chapter 8.14

A Mediator for Biospatial Information Systems

Renato Barrera
Universidad Nacional Autónoma de México, Mexico

Abraham Alcántara
Universidad Nacional Autónoma de México, Mexico

Carlos Alegria
Universidad Nacional Autónoma de México, Mexico

Ana L. Ávila
Universidad Nacional Autónoma de México, Mexico

David Esparza
Universidad Nacional Autónoma de México, Mexico

Augusto D. Hernández
Universidad Nacional Autónoma de México, Mexico

Ivonne Rivera
Universidad Nacional Autónoma de México, Mexico

ABSTRACT

This article presents a system to enable access to those Information Systems at the National Autonomous University of Mexico (UNAM) that are related to Biodiversity and the Environment. The system in question associates existing Geographic Information Systems (GIS’s) as well as standard relational databases in a federation, allows the contents of the individual GIS (or relational databases) to be consulted in a manner transparent to the user, and permits the exports of the underlying systems’ data under the corresponding set of permissions. Our approach is based upon three principles: compliance with international standards, reliance upon Open Source Software in implementation, and usage of servers of proven reliability and robustness.

[Article copies are available for purchase from InfoSci-on-Demand.com]
INTRODUCTION

Mexico occupies a very high position in biodiversity, especially for reptiles, mammals, amphibians, vascular plants, and birds. The National University of Mexico (UNAM) and other institutions possess a treasure trove of collections gathered since the 19th century, and these unique collections continue to grow exponentially. Several research centers at UNAM as well as other Mexican institutions, such as the National Commission for Biodiversity (CONABIO), the National Polytechnic Institute, or U. of the Americas, have developed Geographic Information Systems (GIS) concerning biodiversity and spatial information. Because they have been developed independently, the GIS’s reflect both the idiosyncrasies of the software used in their implementation as well as the peculiarities of each developing institution. A challenge stems from this multiplicity of origins: To allow transparent access to the information of these dissimilar sources to enable their joint exploitation.

To answer this challenge, there has been research on systems’ integration (Levy, 2001; McBrien and Poulavassilis, 2003) and on methods to match ontologies (Aumueller et al., 2005), as well as international standards for specifying:

1. Common schemata for defining the fields for exchanging biodiversity information (Graham et al., 2004).
2. Protocols for accessing, via the Internet, data nodes from a diverse array of biodiversity collections (Graham et al., 2004).
4. Canonical operations for accessing spatial data (Clementini et al., 1993; Worboys, 2004).
6. Query predicates for spatial features in a Web environment (Clementini et al., 1993).
7. The representation of semi-structured spatial information (Lake et al., 2004) by means of a Geographic Markup Language (GML), a standard developed by the Open Geospatial Consortium.

UNAM has established a University Biodiversity System (UNIBIO) that states as its mission: “the generation of a system that allows the capture, organization and public access to Mexico’s biodiversity information”. Parallel to UNIBIO, UNAM has established the UNIGEO project to unify its geo-spatial information using the standards of the Open Geospatial Consortium (OGC) (Lake et al., 2004). The Mediator system described in this article has been developed by UNIBIO to unify it with UNIGEO, using international standards to bridge the gaps between the different data sources.

The remaining sections of this document describe related work, explain the system’s goals and components, present the architecture of the two main constituents of our project to, finally, give results and enumerate conclusions.

REVIEW OF RELATED SYSTEMS

As mentioned in the introduction, the mediator project involves the integration of two types of information systems: the ones concerning biological collections and those related to geo-spatial information. This section describes work in each of these two fields.

Accessing Biological Collections

Two main XML schemata are used for information exchange and data integration on biological collections, namely, the Darwin Core and the schema of the Task Group on Access to Biological Collection Data (ABCD) (Sarkar, 2007; Darwin Core, 2003; ABCD, 2005). These schemata have the following characteristics: