



OpenSearch technology for geospatial resources discovery

Fabrizio Papeschi (1), Boldrini Enrico (1), Paolo Mazzetti (1,2)

(1) Institute of Methodologies for Environmental Analysis of the Italian National Research Council (IMAA-CNR) {papeschi, boldrini, mazzetti}@imaa.cnr.it, (2) University of Florence at Prato, Piazza Ciardi, 25 59100 Prato, ITALY {mazzetti}@imaa.cnr.it

In 2005, the term Web 2.0 has been coined by Tim O'Reilly to describe a quickly growing set of Web-based applications that share a common philosophy of "mutually maximizing collective intelligence and added value for each participant by formalized and dynamic information sharing".

Around this same period, OpenSearch a new Web 2.0 technology, was developed. More properly, OpenSearch is a collection of technologies that allow publishing of search results in a format suitable for syndication and aggregation. It is a way for websites and search engines to publish search results in a standard and accessible format. Due to its strong impact on the way the Web is perceived by users and also due its relevance for businesses, Web 2.0 has attracted the attention of both mass media and the scientific community. This explosive growth in popularity of Web 2.0 technologies like OpenSearch, and practical applications of Service Oriented Architecture (SOA) resulted in an increased interest in similarities, convergence, and a potential synergy of these two concepts. SOA is considered as the philosophy of encapsulating application logic in services with a uniformly defined interface and making these publicly available via discovery mechanisms. Service consumers may then retrieve these services, compose and use them according to their current needs. A great degree of similarity between SOA and Web 2.0 may be leading to a convergence between the two paradigms. They also expose divergent elements, such as the Web 2.0 support to the human interaction in opposition to the typical SOA machine-to-machine interaction.

According to these considerations, the Geospatial Information (GI) domain, is also moving first steps towards a new approach of data publishing and discovering, in particular taking advantage of the OpenSearch technology. A specific GI niche is represented by the OGC Catalog Service for Web (CSW) that is part of the OGC Web Services (OWS) specifications suite, which provides a set of services for discovery, access, and processing of geospatial resources in a SOA framework. GI-cat is a distributed CSW framework implementation developed by the ESSI Lab of the Italian National Research Council (CNR-IMAA) and the University of Florence. It provides brokering and mediation functionalities towards heterogeneous resources and inventories, exposing several standard interfaces for query distribution.

This work focuses on a new GI-cat interface which allows the catalog to be queried according to the OpenSearch syntax specification, thus filling the gap between the SOA architectural design of the CSW and the Web 2.0. At the moment, there is no OGC standard specification about this topic, but an official change request has been proposed in order to enable the OGC catalogues to support OpenSearch queries. In this change request, an OpenSearch extension is proposed providing a standard mechanism to query a resource based on temporal and geographic extents. Two new catalog operations are also proposed, in order to publish a suitable OpenSearch interface. This extended interface is implemented by the modular GI-cat architecture adding a new profiling module called "OpenSearch profiler". Since GI-cat also acts as a clearinghouse catalog, another component called "OpenSearch accessor" is added in order to access OpenSearch compliant services.

An important role in the GI-cat extension, is played by the adopted mapping strategy. Two different kind of mappings are required: query, and response elements mapping. Query mapping is provided in order to fit the simple OpenSearch query syntax to the complex CSW query expressed by the OGC Filter syntax. GI-cat internal data model is based on the ISO-19115 profile, that is more complex than the simple XML syndication formats,

such as RSS 2.0 and Atom 1.0, suggested by OpenSearch. Once response elements are available, in order to be presented, they need to be translated from the GI-cat internal data model, to the above mentioned syndication formats; the mapping processing, is bidirectional. When GI-cat is used to access OpenSearch compliant services, the CSW query must be mapped to the OpenSearch query, and the response elements, must be translated according to the GI-cat internal data model.

As results of such extensions, GI-cat provides a user friendly facade to the complex CSW interface, thus enabling it to be queried, for example, using a browser toolbar.