Semantic geo-catalog: a scenario and requirements

Pavel Shvaiko[1], Lorenzino Vaccari[2], and Gaia Trecarichi[3]

[1] TasLab, Informatica Trentina S.p.A., Trento, Italy
[2] Autonomous Province of Trento, Trento, Italy
[3] DISI, University of Trento, Trento, Italy

Abstract. In this short paper we present a scenario and requirements for ontology matching posed by a geographical application, namely a semantic geo-catalog, which is an integral part of any spatial data infrastructure (SDI). It enables semantic interoperability among various geo-data and geo-service providers, and thus, contributes to the harmonization of geo-information.

Introduction. The need for coherent and contextual use of geographic information between different stakeholders, such as departments in public administrations, formed the basis for a number of initiatives aiming at sharing of spatial information, e.g., the INfrastructure for SPatial InfoRmation in Europe (INSPIRE)[1], see also [8, 10]. In this paper, we focus on a particular component of the INSPIRE architecture, which is a discovery service, that ought to be implemented by means of the Catalogue Service for the Web (CSW)[2], a recommendation of the Open Geospatial Consortium (OGC).

There have been provided several implementations of the CSW-based geo-catalog, e.g., GeoNetwork[3]. A first attempt to provide a semantic geo-catalog has been made in [6], though it was based on a single ontology approach. The approach in [9] proposed an OWL profile for CSW. Finally, the 52°North semantics community[4] proposed to encapsulate ontology repositories by OGC services. In turn, the problem of ontology matching [2] in geo applications has been rarely addressed [7], with some exceptions, such as in [1, 5, 10]. The contribution of this paper includes a specific scenario and requirements for ontology matching posed by a semantic geo-catalog to be realized within the SDI of the Autonomous Province of Trento (PAT).

Scenario. Figure 1 shows a high-level architecture for the semantic geo-catalog system-to-be. Users can issue queries, such as *Trentino mountain hovels reachable with main roads*. The query and search results, such as a map of hovels, are handled by the Trentino geo-portal[5] implemented within the BEA ALUI framework. The geo-catalog will be based on the GeoNetwork open source, while its semantic extension will be designed and developed on top of SWeb[6] search and matching technologies [3, 4]. Specifically, user queries will be analyzed in order to extract concepts out of labels. Then, these are matched at run time against the *universal knowledge* of the SWeb system (SWebUK). In turn, GeoNetwork will contain domain specific ontologies (e.g., Agrovoc) which are associated with geo-metadata and matched with the SWebUK at design time.

[1] http://www.ec-gis.org/inspire/
[2] http://www.opengeospatial.org/standards/cat
[3] http://geonetwork-opensource.org/
[4] http://52north.org
[5] http://www.territorio.provincia.tn.it/
[6] http://www.dit.unitn.it/~knowdive/description.php

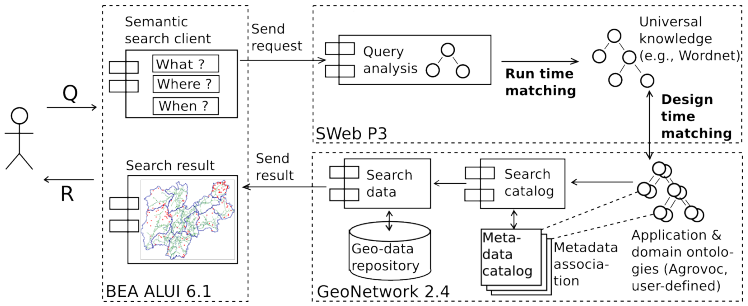

Fig. 1: A high-level architecture for geo-service discovery.

Requirements. There are six general key requirements indicated by INSPIRE, three of which are going to be monitored (for the discovery service), such as: *performance* - to send one metadata record within 3s.; *availability* - service up by 99% of time; *capacity* - 30 simultaneous service requests within 1s. These requirements only put constraints on run time matching needed between the user query and ontologies of the system, that is, the time elapsed between query issue and search results (metadata records) returned should be at most of 3s., etc. Matching results can be approximate here, though their correctness (precision) is preferred over completeness (recall). As for the design time matching between the SWebUK and the domain specific ontologies of the geo-catalog, it can be performed off-line (semi-automatically with sound and complete alignment) when any of these knowledge sources evolves.

Conclusions and future work. In this short paper we have presented a scenario and requirements for ontology matching within a geo-information application, which is a semantic geo-catalog. Future work proceeds at least in the following directions: (i) formalization and in-depth study of the scenario, and (ii) implementation and evaluation of the semantic geo-catalog in order to bring it to production in the SDI of PAT.

Acknowledgments. This work has been supported by the *TasLab network* project funded by the European Social Fund under the act n. 1637 (30.06.2008) of PAT.

References

1. I. Cruz and W. Sunna. Structural alignment methods with applications to geospatial ontologies. *Transactions in Geographic Information Science*, 12(6):683–711, 2008.
2. J. Euzenat and P. Shvaiko. *Ontology matching*. Springer, 2007.
3. F. Giunchiglia, U. Kharkevich, and I. Zaihrayeu. Concept search. In *Proc. of ESWC*, 2009.
4. F. Giunchiglia, F. McNeill, M. Yatskevich, J. Pane, P. Besana, and P. Shvaiko. Approximate structure-preserving semantic matching. In *Proc. of ODBASE*, 2008.
5. K. Janowicz, M. Wilkes, and M. Lutz. Similarity-based information retrieval and its role within spatial data infrastructures. In *Proc. of GIScience*, 2008.
6. P. Maué. An extensible semantic catalogue for geospatial web services. *Journal of Spatial Data Infrastructures Research*, 3:168–191, 2008.
7. P. Shvaiko and J. Euzenat. Ten challenges for ontology matching. In *Proc. of ODBASE*, 2008.
8. P. Smits and A. Friis-Christensen. Resource discovery in a European Spatial Data Infrastructure. *IEEE Transactions on Knowledge and Data Engineering*, 19(1):85–95, 2007.
9. K. Stock, M. Small, Y. Ou, and F. Reitsma. OGC catalogue services - OWL application profile of CSW. Technical report, Open Geospatial Consortium, 2009.
10. L. Vaccari, P. Shvaiko, and M. Marchese. A geo-service semantic integration in spatial data infrastructures. *Journal of Spatial Data Infrastructures Research*, 4:24–51, 2009.