

An Ontology Service for Geographical Content

Eero Hyvönen, Robin Lindroos, Tomi Kauppinen, and Riikka Henriksson

Semantic Computing Research Group (SeCo)
Helsinki University of Technology and University of Helsinki
first.last@tkk.fi, <http://www.seco.tkk.fi>

Abstract. Geographic place names are widely used in databases, but are semantically often highly ambiguous. For example, there are 491 places in Finland sharing the same name “Isosaari” (great island) that are instances of several geographical classes, such as Island, Forest, Field, Peninsula, Inhabited area, etc. Referencing unambiguously to a particular “Isosaari” on the semantic web, either when annotating content or during information retrieval, can be quite problematic and requires usage of advanced search methods and maps for semantic disambiguation. This paper presents an ontology server, ONKI-Paikka, for solving the place finding and place name disambiguation problem. The system currently contains 800,000 place names in Finland and millions of places in other countries. In ONKI-Paikka, places can be found by a faceted search engine, combined with semantic autocompletion and a map service for constraining search and for visualizing results. The service can be connected to applications cost-effectively by using Ajax-technology in the same spirit as Google Maps, used in ONKI-Paikka as a subservice.

1 SUO Place Ontology

ONKI-Paikka¹ is based on the Finnish Place Ontology SUO [2] being developed in the National Semantic Web Ontology Project FinnONTO² (2003-2007) in Finland [1]. The SUO has been populated with 1) place information from the Geographic Names Register (GNR) provided by the National Land Survey of Finland³ and with 2) place information from the GEOnet Names Server (GNS)⁴ maintained by the National Geospatial-Intelligence Agency (NGA) and the U.S. Board on Geographic Names (US BGN). GNR contains about 800,000 names of natural and man-made features in Finland, including data such as place type or feature type and the coordinates of a place. The GNS register contains similar information of about 4,100,000 places around the world excluding places in the United States.

2 ONKI-Paikka Ontology Service

ONKI-Paikka publishes SUO with online ontology services for humans and machines to use. The user interface utilizes AJAX-techniques⁵ for communicating

¹ Operational at <http://demo.seco.tkk.fi/onkipaikka/>

² <http://www.seco.tkk.fi/projects/finnonto/>

³ <http://www.maanmittauslaitos.fi/>

⁴ <http://earth-info.nga.mil/gns/html/>

⁵ <http://ajax.org/>

with the ONKI-Paikka database server containing the place instances. The place finder contains a simple faceted search engine for narrowing the search along the following dimensions: 1) Place name facet filters matching place instances using string autocompletion for their labels. 2) Place type facet of place types (City, Island, Cemetery, etc.) is used to focus search to places of desired types. 3) Language facet limits search to place names in given languages (Finnish, Swedish, English, Lappish, etc.). 4) Time facet is used to focus search on historical place names. 5) Map facet allows the user to specify a polygon area in which to search for the place on the map [2]. The polygon functions as a narrowing criteria for the search in the same way as category selections in the other facets. Using ONKI-Paikka one can e.g. find all places of 1) some type with 2) names beginning with given letters and that are 3) inside a polygon on the map. In this way the end-user can find the intended place reference out of millions of resources more easily.

ONKI-Paikka uses Google Maps API⁶ for visualization, and can be connected to and utilized in legacy systems using AJAX and mash-up techniques with precise references to the underlying URIs in the ontology.

A technical challenge with the ONKI-Paikka service is the large size of the SUO. The response times of the faceted search were several minutes when using a straight-forward Jena⁷ implementation. To solve the efficiency problem, only the SUO ontology classes, without the instances, are stored in a Jena model, and a separate indexing layer was created for the millions of instances. The indexing of the instances is created in a relational database (MySQL) making complex combinational searches possible and relatively fast.

Acknowledgments

Our research is funded by the Finnish Funding Agency for Technology and Innovation (Tekes) and a consortium of 36 companies and public organizations.

References

1. E. Hyvönen, K. Viljanen, E. Mäkelä, T. Kauppinen, T. Ruotsalo, O. Valkeapää, K. Seppälä, O. Suominen, O. Alm, R. Lindroos, T. Käsälä, R. Henriksson, M. Frosterus, J. Tuominen, R. Sinkkilä, and J. Kurki. Elements of a national semantic web infrastructure—case study Finland on the semantic web. In *Proceedings of the First International Semantic Computing Conference (IEEE ICSC 2007)*, Irvine, California, September 2007. IEEE Press, forth-coming.
2. Tomi Kauppinen, Riikka Henriksson, Jari Väättäinen, Christine Deichstetter, and Eero Hyvönen. Ontology-based modeling and visualization of cultural spatio-temporal knowledge. In *Developments in Artificial Intelligence and the Semantic Web - Proceedings of the 12th Finnish AI Conference STeP 2006*, October 26-27 2006.

⁶ <http://maps.google.com>

⁷ <http://jena.sourceforge.net/>

Demonstration

ONKI-Paikka ontology service will be demonstrated at the conference on-line:
<http://www.seco.tkk.fi/services/onkipaikka/>

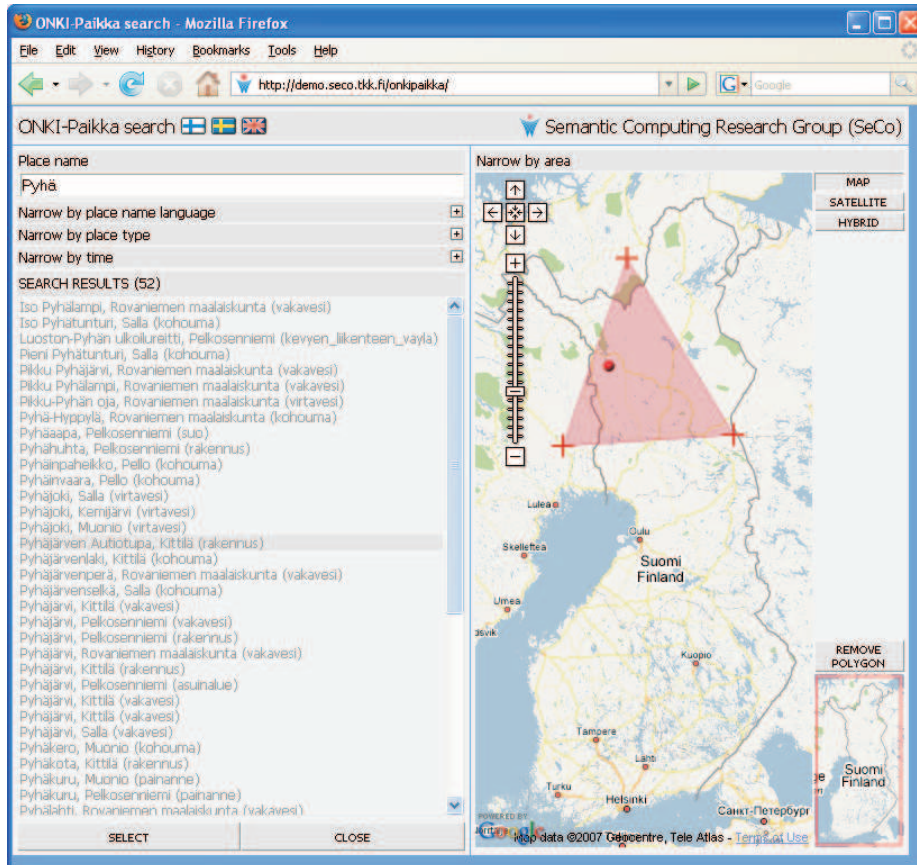


Fig. 1. A view of the ONKI-Paikka. The facets for querying places are on the left, and the map for visualizing search results and for constraining search further by a polygon is on the right.