

Eero Hyvönen

Esko Ikkala, Eero Hyvönen, Jouni Tuominen

Aalto University and University of Helsinki

An Ontology of World War II Places for Linking and Enriching Heterogeneous Historical Data Sources

This paper presents a historical geo-ontology of some 34 000 historical places, based on old maps of the war zone between Finland and the Soviet Union during the World War II. This area was annexed to the Soviet Union after the war, and modern maps have only Soviet or Russian names, making it impossible to use modern gazetteers to describe primary source data of the war, such as photographs, articles, war diaries etc., in which original Finnish placenames are used. As a use case, the ontology was applied to linking and enriching data in nine applications of the WarSampo semantic portal (<http://sotasampo.fi/en>) that have been used by tens of thousands of end users. The paper overviews lessons learned in this work. The ontology was constructed using geocoding data and existing Finnish gazetteers, covering the whole area of Finland and beyond. Old maps of the war zone were rectified and linked into the system via an old map service based on Map Warper (<https://github.com/timwaters/mapwarper>), developed at the Public Library of New York. Named entity linking of geonames was used to linking 160 000 photo captions, over 1000 event descriptions, and 3400 magazine articles. The resulting data is available as 5-star Linked Open Data at the Linked Data Finland service (<http://ldf.fi>), with content negotiation, a SPARQL endpoint, and additional services for reusing the data.

Keywords: ontologies, linked data, geography, semantic web