

Toward next generation of gazetteer: utilizing GeoSPARQL for developing Linked Geoname Data

Dong-Po Deng^{1,3*}, Tyng-Ruey Chuang^{1,2}

1 Institute of Information Science, Academia Sinica

2 Research Center for Information Technology Innovation, Academia Sinica

3 Department of GeoInformation Process, ITC, Twente University

E-mail: dongpo@iis.sinica.edu.tw

Position technologies such as GPS are directly geo-referenced with geographic coordinates, however, the use of geographic names for indicating locations is instinctive. The value of geographic names is not only used for geographic reference but also represents the landscape of culture and social. Thus, the concepts of geographic names are often complicated, diversified, ambiguous and multi-scaled geospatial objects. There is a need to specify the place name to canonical and interchangeable geospatial knowledge.

Linked data is a new research area which studies how to make data available on the Web, and to interconnect data with the aim of increasing its value for users. Each entry representing a fact in LOD datasets has a Unique Resource Identifier (URI) which is referenceable and linkable on the Web. The high interconnectivity between entries potentially increases discoverability, reusability, and the utility of information. Therefore, if geographic names can be identified and connected to entries of LOD, the semantics of geographic names would be disambiguated, so that the UGC could be easier to process.

The aim of this study is to enhance the semantics of geographic names through the use of LOD, so that the data of geographic names can be reusable and discoverable. This paper describes our experiences in developing a Linking Open Data (LOD) of Taiwanese geographic names (LOD TGN). By using standard vocabularies, we develop an ontology for formalizing the concepts of geographic names in spatial, temporal, and thematic components. Based on the ontology, we transfer the entries of gazetteers to Resources Description Framework (RDF) as well as interlink geographic names to outbound linked data resources such historical geonames and geonames.org. Moreover, a SPAQL endpoint is developed for serving spatiotemporal SPARQL queries.