

Noisy Semantic Data Processing in Seoul Road Sign Management System

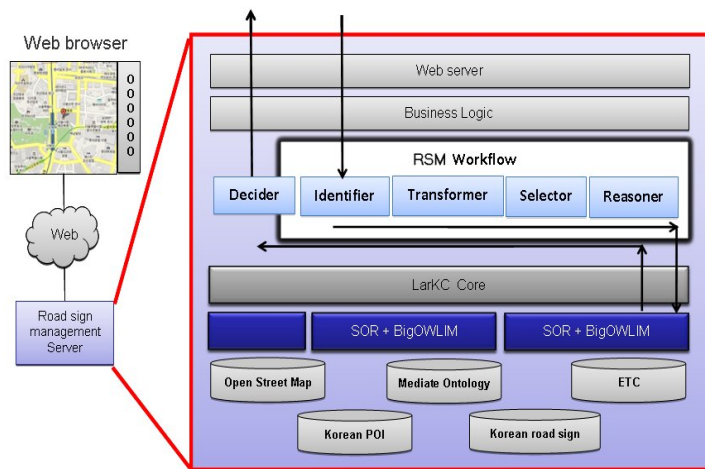
Zhisheng Huang¹, Jun Fang², Stanley Park³, and Tony Lee³

1. Department of Computer Science, Vrije Universiteit Amsterdam, The Netherlands

2. School of Automation, Northwestern Polytechnical University, China

3. Saltlux Inc., Seoul, Korea

The Seoul Road Sign Management (RSM) is a system which provides the semantic integration of LOD's Linked Geo Data and Open Street Map with Korean POI data set. That is an attempt to develop intelligent road sign management system based on the LarKC platform. The RSM data set contains over 1.1 billion triples of semantic data. However, significant amount of the RSM data are noisy (e.g., inconsistent, partial, or erroneous). We have facilitated the RSM system with the capability of processing and reasoning with noisy semantic data, so that the RSM system is robust enough to return intended answers in spite of the poor quality of the semantic data.



Dataset	Features and Operations
Linked Geo Data (LGD)	1 billion triples in WGS84 coordinate Loading LGD full and extracted POIs written in English
Open Street Map (OSM)	Extracting all way information in WGS84 coordinate Selecting and importing 2 million triples for Seoul
POI Data in Korea (KPOI)	1 million POIs related with road signs written in Korean Around 4 million triples
Seoul road sign data (RSD)	Diverse data set of Seoul road signs Extracting half million triples from RDB Converting TM coordinate into WGS84 coordinate
Korean road sign regulations (RSR)	Around 30 Regulations of road sign about positioning and naming. Changing into SparQL for validation check
Mediate Ontology (MO)	Ontology linking between OSM, KPOI and RSD or other data Expressivity : subClassOf, subPropertyOf, sameAs, inverseOf

Table 1. RSM data set

