Centrality-based Selection of Semantic Resources for Geosciences

Otakar Cerba (1) and Karel Jedlicka (2)

(1) Department of Geomatics, University of West Bohemia, Plzen, Czech Republic (cerba@kgm.zcu.cz), (2) Department of Geomatics, University of West Bohemia, Plzen, Czech Republic (smrcek@kgm.zcu.cz)

Semantical questions intervene almost in all disciplines dealing with geographic data and information, because relevant semantics is crucial for any way of communication and interaction among humans as well as among machines. But the existence of such a large number of different semantic resources (such as various thesauri, controlled vocabularies, knowledge bases or ontologies) makes the process of semantics implementation much more difficult and complicates the use of the advantages of semantics. This is because in many cases users are not able to find the most suitable resource for their purposes.

The research presented in this paper introduces a methodology consisting of an analysis of identical relations in Linked Data space, which covers a majority of semantic resources, to find a suitable resource of semantic information.

Identical links interconnect representations of an object or a concept in various semantic resources. Therefore this type of relations is considered to be crucial from the view of Linked Data, because these links provide new additional information, including various views on one concept based on different cultural or regional aspects (so-called social role of Linked Data).

For these reasons it is possible to declare that one reasonable criterion for feasible semantic resources for almost all domains, including geosciences, is their position in a network of interconnected semantic resources and level of linking to other knowledge bases and similar products.

The presented methodology is based on searching of mutual connections between various instances of one concept using "follow your nose" approach. The extracted data on interconnections between semantic resources are arranged to directed graphs and processed by various metrics patterned on centrality computing (degree, closeness or betweenness centrality).

Semantic resources recommended by the research could be used for providing semantically described keywords for metadata records or as names of items in data models. Such an approach enables much more efficient data harmonization, integration, sharing and exploitation.

* * * *